



UNIVERSITÀ DEGLI STUDI DI PALERMO

Dipartimento di Biomedicina Sperimentale e Neuroscienze Cliniche (BioNeC)
Dottorato di ricerca in Oncobiologia Sperimentale

Coordinatore Prof.ssa Renza Vento

Ricerca di biomarcatori nel carcinoma epatocellulare

BIO/19 - Genetica

Candidato

Dr. Daniele Balasus

Tutor

Prof.ssa Michela Giuliano

Co-Tutor

Prof. Mario La Farina

Il Coordinatore del Dottorato

Prof.ssa Renza Vento

Triennio 2011-2013-Ciclo XXIV

Indice

Introduzione	4
1. Il carcinoma epatocellulare	5
1.1. Epidemiologia del carcinoma epatocellulare	5
1.2. Eziologia del carcinoma epatocellulare	8
1.3. Genetica del carcinoma epatocellulare	10
2. I polimorfismi a singolo nucleotide (SNPs) e il carcinoma epatocellulare	13
2.1. L'SNP rs231775 A/G sul gene <i>CTLA4</i>	14
2.2. L'rs2304052 T/C sul gene <i>SPARC</i>	15
2.3. L'rs4444903 G/A sul gene <i>EGF</i>	16
2.4. L'rs430397 A/G sul gene <i>GRP78</i>	17
2.5. L'rs738409 G/C sul gene <i>PNPLA3</i>	18
3. Epigenetica e metiloma nel carcinoma epatocellulare	19
Obiettivi	22
Risultati	24
1. Selezione dei polimorfismi a singolo nucleotide	25
2. Raccolta, selezione dei campioni e progettazione dello studio	25
3. Genotipizzazione mediante PCR-RFLP	26
4. Studio caso-controllo: analisi dei dati	28
5. Genotipizzazione KASPar™	32
6. Valutazione dei livelli di mRNA di <i>GRP78</i>	35
7. Analisi dello stato di metilazione del promotore di <i>GRP78</i>	37
8. Valutazione dell'espressione di <i>GRP78</i>	40
Pazienti, materiali e metodi	42
1. Casi di carcinoma, di cirrosi epatica e controlli	43

2. Polimorfismi a singolo nucleotide	43
3. Estrazione e valutazione della qualità degli acidi nucleici	44
4. PCR-RFLP	45
5. Colture cellulari	48
6. Genotipizzazione Kaspar™	48
7. Real Time PCR quantitativa	49
8. MeDip-qPCR (Methylated DNA immunoprecipitation-qPCR)	50
9. Analisi Statistica	53
10. Immunoistochimica	53
11. Soluzioni	54
 Discussione	 57
 Conclusioni	 67
 Bibliografia	 79
 Ringraziamenti	 77

Introduzione

1. Il carcinoma epatocellulare

Il Carcinoma epatocellulare (CE) è responsabile dell'85-90% circa delle neoplasie a carico del fegato, è il quinto più comune tipo di cancro ed è la terza causa di morte per cancro nel mondo. Purtroppo ci si aspetta un peggioramento di questi indici nei prossimi anni. Questo tumore è associato spesso ad uno stato di cirrosi epatica, che deriva da un'inflammatione cronica del tessuto epatico.

Il CE ha delle singolari caratteristiche epidemiologiche ed eziologiche. La sua incidenza e mortalità infatti variano parecchio in base alla localizzazione geografica, ai gruppi etnici, alle razze e al periodo temporale in cui le si analizza. Il più alto tasso di mortalità si registra nelle regioni dell'Africa sub-Sahariana e dell'Asia dell'Est (90% circa), questo dato si ridimensiona per la maggior parte dei paesi industrializzati (fig. 1). Nonostante oggi siano a disposizione una grande quantità di dati riguardanti l'epidemiologia dell'CE, gli studi che li hanno prodotti presentano diversi limiti, soprattutto perché hanno utilizzato diverse metodologie e sono stati effettuati in diversi momenti. Per questo è necessario un nuovo approccio globale, che possa essere utile alla prevenzione e al trattamento della patologia.

La variabilità nell'incidenza e mortalità si intreccia, ed è fortemente correlata, con l'eziologia della patologia. I fattori di rischio per il CE infatti sono tanti e sono diversamente distribuiti nelle aree geografiche e tra le differenti etnie.

E' ormai assodato che il CE si sviluppi per quasi la totalità dei casi in un tessuto epatico non normale. Perlopiù infatti (dal 70% al 90% dei casi) il CE segue ad uno stato di cirrosi epatica o di inflammatione cronica (Delhay M., *et al.*, 1996).

1.1 Epidemiologia del carcinoma epatocellulare

Come già accennato l'incidenza e la mortalità del CE sono strettamente dipendenti dalla localizzazione geografica. Le maggiori fonti di variabilità sono legate alla razza o etnia, al sesso, all'età e alla differente distribuzione dei fattori di rischio.

I soggetti di sesso maschile hanno un rischio maggiore di CE rispetto a quelli di sesso femminile, con un rapporto uomo:donna che va da 2:1 a 4:1. In alcuni paesi si raggiunge anche un rapporto di 5:1 (Italia), ma in generale in tutta l'Europa, specialmente del nord, questo rapporto si aggira intorno a 4:1. Nelle popolazioni invece a più alto rischio si arriva anche ad un rapporto di 2,4:1 (El Serag HB., *et al.*, 2007). Le ragioni che portano a questa differente suscettibilità "sesso-specifica" sono soprattutto legate all'esposizione ai fattori di rischio. I soggetti di sesso maschile sono infatti più esposti all'infezione da virus dell'epatite C e B, al consumo di alcool e al fumo di sigaretta. Molte

ricerche tuttavia supportano l'ipotesi di un'influenza degli androgeni che agirebbero da fattore protettivo e del testosterone che invece sarebbe un ulteriore fattore di rischio (El Serag HB., *et al.*, 2007).

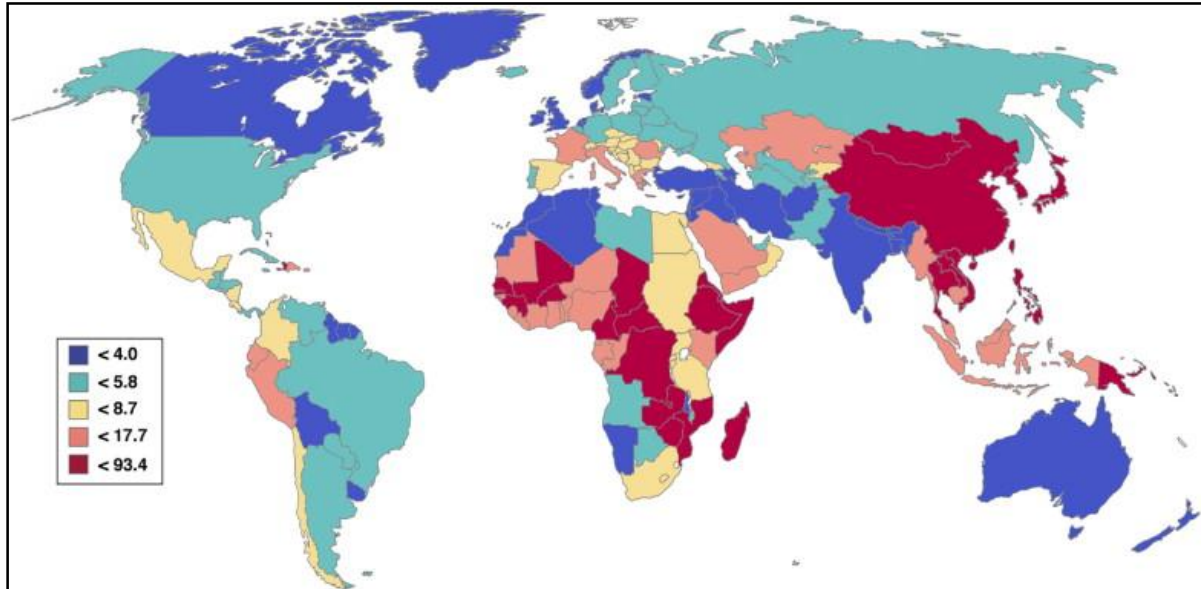


Figura 1: Variabilità nella mortalità globale per HCC (El Serag HB., *et al.*, 2007).

I fattori di rischio maggiori per il CE variano parecchio nel mondo. Nelle aree a maggiore incidenza il fattore di rischio principale è l'infezione da HBV. Spesso infatti in queste aree il virus si trasmette per via materno-fetale. Un esempio concreto è lo studio sul cancro al fegato nel Gambia. Da questa ricerca si evince infatti che il 60% dei casi di HCC su un campione di 216 è HBV positivo (Venook AP., *et al.*, 2010).

Al contrario del resto dei paesi asiatici, in Giappone è l'HCV il virus più diffuso da epatite, l'infezione cronica causa, in alcuni casi, cirrosi epatica, quindi CE. Nei paesi invece a bassa incidenza, i fattori di rischio si combinano in maniera differente (El Serag HB., *et al.*, 2007). In alcuni di essi è l'infezione da HBV, in altri quella da HCV ad essere il principale fattore di rischio. Nei paesi dove l'infezione da HBV non è endemica, l'infezione da HCV e la cirrosi alcolica sono considerati i maggiori fattori di rischio (fig. 2). Nella popolazione in analisi nel presente studio, ovvero quella siciliana, il maggiore fattore di rischio è l'infezione da HCV. Da un recente studio (Soresi M., *et al.*, 2010) infatti è chiaro che la maggioranza dei casi di CE (80% circa) è dovuta ad un'infezione cronica da HCV. Prima del 1990 non esisteva un test di laboratorio capace di rilevare la presenza dell'HCV nel sangue infetto. Nel periodo precedente infatti si è registrato un incremento delle infezioni e di conseguenza un'impennata dei casi di cirrosi epatica e di CE, dovuta alla contaminazione delle scorte di sangue

nei presidi ospedalieri dell'Europa del nord e del sud. Questa è una, ma sicuramente non l'unica, tra le spiegazioni possibili, per la singolare distribuzione dei fattori di rischio nella popolazione siciliana.

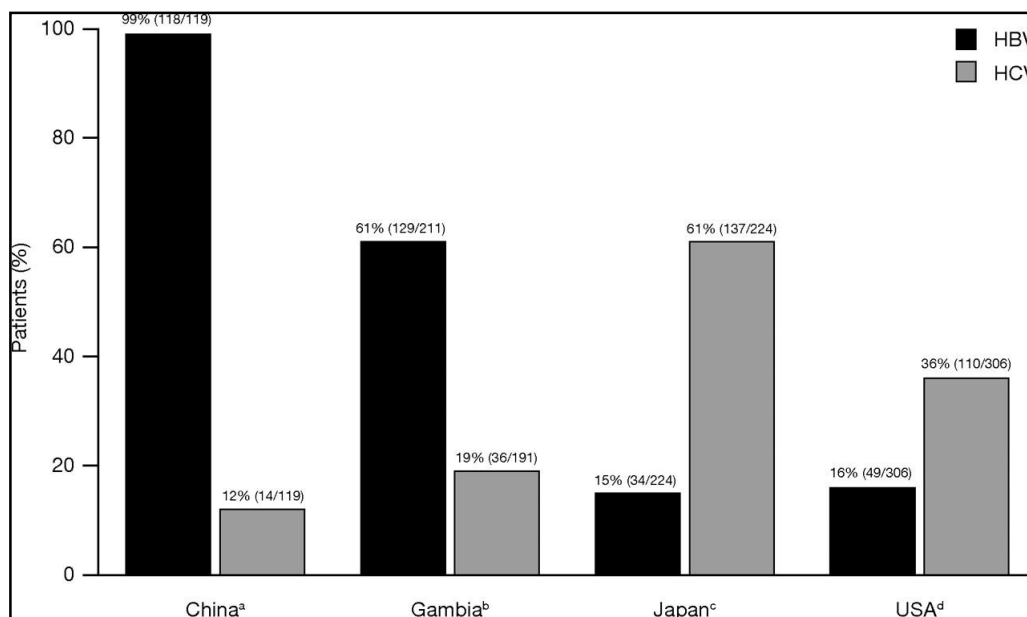


Figura 2: variazioni nella prevalenza dell'infezione da HBV ed HCV (Venook AP., *et al.*, 2010).

Un altro fattore di variabilità è il genotipo virale dell'HBV. Ad oggi sono stati classificati otto genotipi, da A a H. I pazienti infettati con il genotipo C hanno un maggiore rischio di sviluppare il CE. Anche l'infezione da HIV (virus dell'immunodeficienza umana) incrementa il rischio di CE nei soggetti HBV positivi (Venook AP., *et al.*, 2010). Questo dato è di particolare rilevanza e spiega in parte l'elevata incidenza di CE nell'Africa sub-Sahariana, dove il virus dell'HIV è molto diffuso.

Un altro fattore di rischio endemico, in Asia e in Africa Sub-Sahariana, è il consumo di cibi contaminati con aflatossina B1, un potente epatocarcinogeno e un potenziale contaminante di legumi, grano, noccioline ecc... (Venook AP., *et al.*, 2010).

Negli ultimi due decenni in alcuni paesi dell'America e in Europa del sud si è registrato un incremento dei casi di CE. Si pensa che questo sia una conseguenza dell'incremento delle infezioni da HCV dovute ai crescenti flussi migratori dai paesi in via di sviluppo e dal crescente consumo di alcool. In altri paesi come il Canada invece è l'obesità uno dei fattori di rischio che ha contribuito all'aumento dei casi di CE (Dyer Z., *et al.*, 2005). L'incidenza del CE tuttavia non è in aumento dovunque. Ad Osaka, in Giappone, ad esempio, l'incidenza è diminuita da 41,9 abitanti su 100.000 a 24 dal 1995 al 2003 (Tanaka H., *et al.*, 2008).

Nel campo dell'oncologia sono stati scoperti spesso dei fattori di rischio, come il fumo di sigarette per il cancro al polmone o l'uso di contraccettivi orali per il cancro al seno, solo studiando la

patologia in una specifica area geografica. Allo stesso modo è fondamentale lo studio in ciascuna popolazione per il CE, una patologia con una grande variabilità di fattori di rischio, le cui concause si intrecciano formando un complicato quadro, differente per ciascuna area geografica del mondo.

La popolazione in analisi nella presente ricerca potrebbe ospitare degli isolati genetici. I genotipi con queste caratteristiche infatti sono peculiari di alcune nicchie di popolazioni che vivono in isolamento geografico e culturale da diverse generazioni. Queste sub-popolazioni derivano interamente da un piccolo numero di individui, grazie ad un effetto “collo di bottiglia”. Qualsiasi ricerca in popolazioni di questo tipo potrebbe dare dei risultati inattesi ma utili a progettare piani di prevenzione o di terapie personalizzate al livello locale.

1.2 Eziologia del carcinoma epatocellulare

Il 70%-90% dei pazienti affetti da CE possiedono un quadro di epatite cronica e/o cirrosi epatica. I maggiori fattori di rischio per il CE sono l'infezione da virus dell'epatite B e C, l'alcool e lo stato denominato NASH (nonalcoholic steatohepatitis). I fattori di rischio minori sono invece il consumo di cibi contaminati con aflatossina B, il diabete, l'obesità e alcune condizioni ereditarie come l'emocromatosi e altri disordini metabolici (Sanyal AJ., *et al.*, 2010).

Tutti quei fattori che contribuiscono a creare le condizioni per la cirrosi epatica patologia sono dei fattori di rischio per il CE (Sanyal AJ., *et al.*, 2010). Un tessuto epatico cirrotico è caratterizzato da una scarsa proliferazione cellulare epatocitaria, a favore invece di una maggiore presenza di fibre e matrice extracellulare, rispetto ad un fegato normale. Un clone epatocitario con un tasso proliferativo deregolato trova quindi delle condizioni favorevoli per la sua espansione, a differenza che in un fegato normale e proliferante. A tutto questo si aggiunge una riduzione della capacità rigenerativa del fegato. Nella cirrosi epatica si riscontra inoltre una maggiore proliferazione delle cellule stellate. Queste cellule contribuiscono ad un'aumentata produzione di proteine della matrice extracellulare, citochine, fattori di crescita, e prodotti dello stress ossidativo. Tutti questi fattori producono un “microambiente” ottimale per lo sviluppo della neoplasia.

Il 50% dei casi di CE, al livello mondiale, sono associati ad un'infezione da HBV (Sanyal AJ., *et al.*, 2010). L'HBV è un virus a DNA a doppio filamento della famiglia degli *Hepadnaviridae*. L'infezione può portare allo sviluppo di CE attraverso meccanismi diretti e indiretti. Per prima cosa causa infiammazione cronica e danno epatocitario, con conseguente incremento della proliferazione epatocitaria, fibrosi e cirrosi. Questi stati favoriscono notoriamente l'espansione di cloni cellulari che hanno accumulato mutazioni su oncogeni, oncosoppressori e altre alterazioni genetiche ed epigenetiche. L'HBV può comunque favorire il CE in maniera diretta. Il virus infatti può integrarsi nel

DNA degli epatociti dell'ospite causando riarrangiamenti cromosomali e incrementando l'instabilità genomica. La proteina regolatoria HBx, inoltre, prodotta dal genoma virale, ha la capacità di transattivare geni coinvolti nel controllo della proliferazione cellulare, con l'attivazione di vie come quelle della protein chinasi C e/o nfκB. Questo può portare ad una deregolazione del ciclo cellulare e a un'interferenza nei processi apoptotici e di riparo del DNA (Sanyal AJ., *et al.*, 2010).

L'HCV, appartiene alla famiglia dei *Flaviviridae* ed ha un genoma a RNA. Esso è incapace di integrarsi nel genoma ospite, infatti sono perlopiù le proteine virali e la risposta immunitaria, da questo indotta, i principali responsabili del processo oncogenico (Jeong SW., *et al.*, 2012). La proteina del capside virale C è probabilmente coinvolta in fenomeni come l'apoptosi, la trasduzione del segnale, la formazione di specie reattive dell'ossigeno (ROS), il metabolismo lipidico, l'attivazione della trascrizione e l'immunomodulazione. La proteina C infatti interagisce con oncosoppressori quali p53, p73 e pRb, con la chinasi ciclino-dipendente p21^{WAF1} e influenza l'attivazione della via delle chinasi Raf/MAPK, la via Wnt/βcatenina e quella del TGFβ. (Jeong SW., *et al.*, 2012). Anche la proteina virale NS5 ha un'attività pro-carcinogenetica. Essa è essenziale per la replicazione virale ed è coinvolta nell'apoptosi, nella trasduzione del segnale, nella produzione di ROS e nel metabolismo lipidico. Infine sono stati individuati diversi SNPs associati al CE su soggetti HCV positivi (Jeong SW., *et al.*, 2012).

L'eccessivo consumo di alcool provoca infiammazione cronica al livello epatico e cirrosi epatica. Non è tuttavia ancora chiaro se l'alcool abbia un effetto cancerogenico diretto. L'alcool è inoltre un importante cofattore nel CE che si sviluppa su soggetti che sono anche HCV e/o HBV positivi, svolgendo così un effetto sinergistico (El Serag HB., *et al.*, 2007). Il meccanismo tramite il quale l'alcool peggiora le condizioni dei soggetti con un'infezione cronica da HCV non è ancora del tutto chiaro. Alcune ipotesi sono: l'incremento del tasso replicativo del virus in presenza di alcool, l'inibizione dell'espressione di Bcl-2 al livello epatico, con la conseguente induzione dell'apoptosi e aumento del danno epatico e infine un aumento del danno ossidativo (Sanyal AJ., *et al.*, 2010).

La patologia denominata NAFLD è molto comune nei paesi occidentali, con il 20% circa degli individui affetti (Sanyal AJ., *et al.*, 2010). Questa si sviluppa in assenza di alcool. Sembra che alcune delle condizioni patologiche che portano allo stato di NAFLD siano l'obesità, il diabete mellito di tipo 2 e la sindrome metabolica. Anche se i meccanismi non sono del tutto chiari è ormai evidente che soprattutto l'insulino resistenza (IR) conduce alla NAFLD. L'IR infatti causa l'accumulo dei trigliceridi al livello epatico, incrementa i livelli di acidi grassi intraepatici, e gioca un ruolo importante nello sviluppo della steatoepatite.

Tra i fattori di rischio minori per il CE c'è l'aflatossina B1. Presente nei cibi contaminati dal fungo *Aspergillus*, l'aflatossina è una micotossina che una volta ingerita viene trasformata nel suo metabolita biologicamente attivo, l'AFB1-exo-8,9-epossido. Questo metabolita si lega al DNA causando dei danni (Sanyal AJ., *et al.*, 2010).

Alcuni studi hanno evidenziato un moderato rischio di CE anche per i soggetti affetti da emocromatosi (El Serag HB., *et al.*, 2007). Altri fattori di rischio lievi per il CE sono la cirrosi biliare primitiva, il deficit da α -1antitripsina, la sifilide e la pericardite. Il consumo di caffè invece sembra essere un fattore protettivo (El Serag HB., *et al.*, 2007).

Almeno 20 SNPs sono stati associati al CE in differenti popolazioni e con una diversa forza. Ciascuno di essi va considerato come fattore di rischio più o meno consistente, limitatamente alle regioni geografiche in cui è stato studiato (Jeong SW., *et al.*, 2012; Sur I., *et al.*, 2013). Si rimanda questo argomento ai capitoli successivi.

Dal quadro appena disegnato appare chiaro che è possibile tenere sotto controllo i soggetti a rischio di CE. Non sempre però è possibile fare uno screening accurato della popolazione a rischio di CE soprattutto per i costi che tale procedura avrebbe. Si stima che nel regno unito il 40% dei pazienti appaiono alla diagnosi come soggetti con il CE come prima patologia epatica. Al contrario in Giappone l'80% dei CE viene diagnosticato ad uno stadio presintomatico (Sanyal AJ., *et al.*, 2010). Questi dati danno un'idea dell'importanza che la prevenzione e la medico sorveglianza possano avere per una riduzione della mortalità del CE e di conseguenza dei costi sanitari. E' necessario quindi ricercare dei marcatori di CE che consentano di individuare e tenere sotto controllo i soggetti a rischio maggiore di CE.

1.3 Genetica del carcinoma epatocellulare

Oltre ai fattori di rischio ambientali, come ampiamente documentato fino ad ora, esistono alcuni fattori di rischio genetici ed ereditari. Possiamo suddividere tali fattori in monogenici e poligenici.

Tra i fattori monogenici c'è il deficit da α 1-antitripsina. La mutazione in questo caso risiede sul gene SERPINA1 sul cromosoma 14. Questo gene codifica per un inibitore di serina proteasi, che è sintetizzata ad alti livelli nel fegato. Entrambi gli alleli sono espressi, quindi un soggetto che eredita un allele normale è parzialmente protetto contro una deficienza severa (Dragani TA., *et al.*, 2010).

Uno dei fattori di rischio poligenici è l'epatite autoimmune. Una rara patologia, le cui cause sono ancora sconosciute e che porta alla progressiva distruzione del parenchima epatico. A parte alcuni SNPs sul locus HLA, nessun altro locus genico in particolare è stato messo in relazione con questa malattia (Dragani TA., *et al.*, 2010). Un'altra patologia poligenica è il diabete. In questo caso si

ipotizza che l'incremento della produzione di IGF (insulin-like growth factor) possa promuovere l'epatocarcinogenesi. Anche l'ipotiroidismo è considerata una patologia poligenica che aumenta il rischio di CE. In particolare sembra centrale il ruolo degli ormoni tiroidei nella mobilizzazione dei lipidi, nella loro degradazione e nell'ossidazione degli acidi grassi. La NASH è un'altra condizione patologica poligenica che è ormai chiaramente associata al CE. In questo caso ci sono tanti SNPs con i quali la NASH è stata messa in relazione (Dragani TA., *et al.*, 2010). Tra gli altri SNPs, possono essere raggruppati in un'unica categoria quelli che aumentano la suscettibilità ad agenti chimici e sostanze tossiche, come l'alcool e l'aflatossina B1, notoriamente processate dal fegato. Alcuni esempi sono gli SNPs sui geni che codificano per i citocromi CYP1A2 e CYP2E1, l'alcool deidrogenasi (ADH), l'aldeide deidrogenasi 2 (ALDH2) e la mieloperossidasi (Dragani TA., *et al.*, 2010). Infine in alcuni paesi, come la Svezia, i fattori ambientali hanno un impatto basso sul rischio di CE, se comparato a quello di molte altre nazioni. In questi paesi infatti sono più comuni i fattori di rischio poligenici familiari (Dragani TA., *et al.*, 2010).

Nonostante oggi ci sia un'ampia conoscenza delle alterazioni genetiche nel CE, non è stato ancora sviluppato un modello preciso e condiviso di epatocarcinogenesi. Già durante lo stadio di cirrosi epatica o nell'epatopatia cronica in generale, si riscontrano diverse anomalie. Ad esempio un accelerato accorciamento dei telomeri provoca l'accumulo di epatociti senescenti e una volta raggiunto un punto critico di accorciamento si instaurano alterazioni molecolari riguardanti l'arresto del ciclo cellulare e l'apoptosi. L'accorciamento telomerico permane per tutta l'epatocarcinogenesi, fino al CE conclamato (Plentz RR., *et al.*, 2007). La lunghezza dei telomeri è mantenuta dall'enzima telomerasi. L'alterazione di quest'enzima è frequente nel CE (90% circa dei casi) (Shiraha H., *et al.*, 2013). Lo stadio di cirrosi è ormai considerato un vero e proprio stadio pre-tumorale. Le alterazioni, in un tessuto epatico cirrotico, provocano, tra le prime anomalie, un'aumentata espressione di p21 e p16, che inducono l'arresto del ciclo cellulare e contribuiscono alla senescenza cellulare (Aravalli RN., *et al.*, 2008). L'iperattivazione delle cellule stellate, altra caratteristica della cirrosi epatica, incrementa la loro stessa sensibilità a fattori di crescita come il PDGF (platelet derived growth factor) e il TGFβ (transforming growth factor beta) (Jeong SW., *et al.*, 2012). In un contesto di infiammazione cronica, come quello cirrotico, inoltre, si osservano alterazioni indotte da molecole pro-infiammatorie come il TNFα. Una di queste è l'attivazione del fattore nucleare κB (nfκB) (El Serag HB., *et al.*, 2007). Questi sembrano essere i primi stadi che portano alla formazione del clone maligno.

Le anomalie molecolari più comuni in una neoplasia sono quelle su p53 ed Rb. Nel CE invece queste due alterazioni sono meno diffuse. Solo il 45% circa dei CE su soggetti HBV positivi e il 13% su HCV positivi infatti presentano delle alterazioni di p53 (Shiraha H., *et al.*, 2013).

Il gene *Rb* si ritrova mutato solo nel 15% dei CE. Nonostante questo, la via di Rb è alterata nell'80% degli epatocarcinomi e in questo caso la repressione di *p16* rappresenta l'anomalia più frequente (Edamoto Y., *et al.*, 2003). PTEN (phosphatase and tensin homolog) è un importante oncosoppressore che regola negativamente la via di regolazione della sopravvivenza cellulare Akt/PI3K. L'assenza o la netta riduzione di PTEN si trova nel 40% dei CE (Edamoto Y., *et al.*, 2003).

Il ruolo degli oncogeni nel CE sembra essere meno importante rispetto a quello degli oncosoppressori. Uno di questi è *Myc*. Il coinvolgimento dell'oncogene *Myc* nel CE è stato ampiamente documentato (Shiraha H., *et al.*, 2013). *Myc* media la risposta cellulare ad alcuni fattori di crescita ed è coinvolto nel metabolismo degli acidi nucleici.

Le vie di segnalazione alterate nel CE sono diverse. Spesso si riattiva la via embrionale di Wnt/ β -catenina. La sua attivazione è dovuta ad alterazioni epigenetiche sull'oncosoppressore *APC* (Shiraha H., *et al.*, 2013) e alle frequenti mutazioni sulla β -catenina, come quelle sull'esone 3 (Lèvy L., *et al.*, 2002). Anche la via embrionale precoce di Hedgehog è coinvolta nell'epatocarcinogenesi (Sicklick JK., *et al.*, 2006). Sono iperattivate le vie di Ras-MAP chinasi, PI3-chinasi e quella attivata da IGF-2 (insulin growth factor-2). Quest'ultima si attiva probabilmente a causa di mutazioni o perdita di eterozigotà sul gene che codifica per il recettore di IGF-2, come si riscontra nel 25-55% dei CE.

Da alcune analisi di trascrittomica nel CE è emersa l'aumentata espressione di alcuni geni, tra questi: TERT, Glypican3, Gankyrina, E-caderina e HGF. Alcuni di questi sono i candidati per una possibile diagnosi precoce di CE (Llovet JM., *et al.*, 2006).

Tra le alterazioni scoperte di recente nel CE ci sono quelle legate ai *microRNA*. I *microRNA* Più frequentemente deregolati nel CE sono miR-21, miR-34a, miR-221/222 e miR-224. Altri come il miR-18 e il miR-20 sono stati associati al grado di differenziazione istologica (Ladeiro Y., *et al.*, 2008).

Un ultimo interessante aspetto è quello legato ai soggetti HCV e HBV positivi. Essi sono caratterizzati infatti da una differente impronta genetica e proteomica. Nei CE legati ad epatiti croniche B sono alterati perlopiù i geni per l'induzione dell'apoptosi, l'arresto del ciclo cellulare e la degradazione della matrice cellulare. Per quanto riguarda invece i CE su soggetti HCV positivi, sono iperespressi ad esempio i geni con effetto antiaoptotico, per l'accelerazione del ciclo cellulare e per la produzione di matrice extracellulare.

Appare quindi chiaro che lo sviluppo del CE è associato a diverse alterazioni che diventano stabili e si accumulano nel corso del tempo, anche se non si sa ancora bene con che ordine. Nella figura 3 si

è provato a schematizzare il processo che porta al CE, accompagnato da alcune ipotesi sulle alterazioni molecolari.

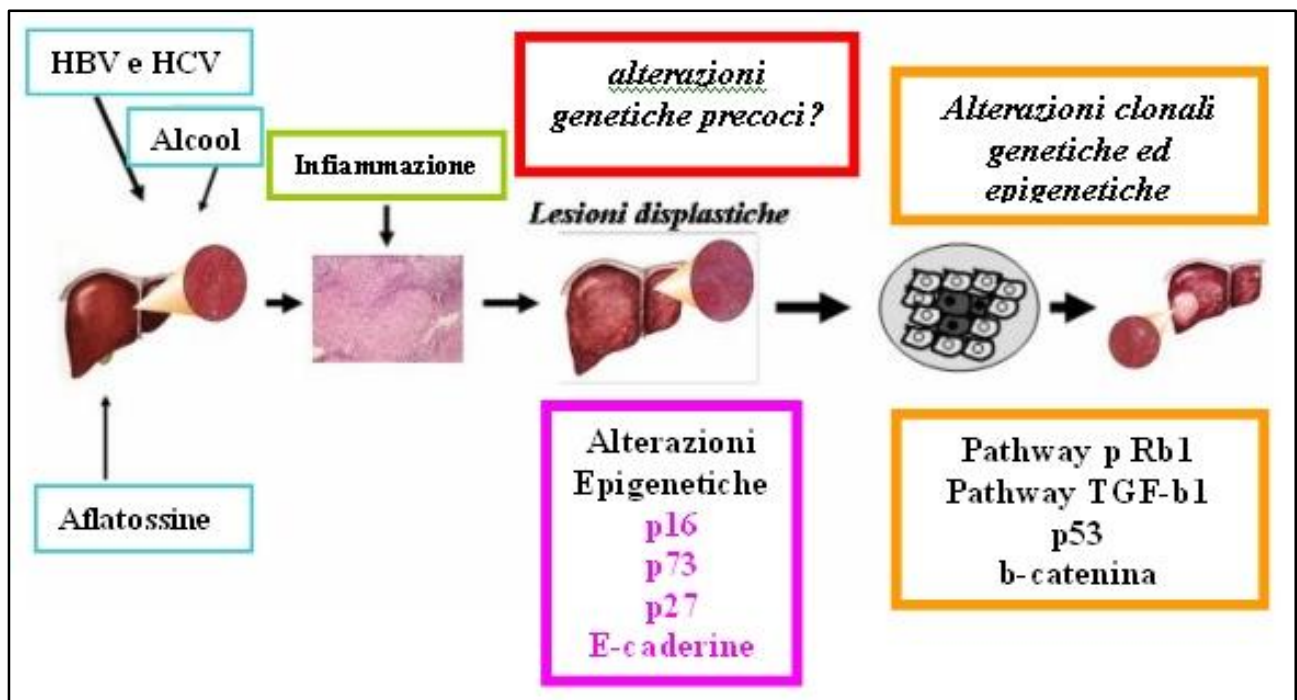


Figura 3: Il processo di epatocarcinogenesi.

2. I polimorfismi a singolo nucleotide (SNPs) e il carcinoma epatocellulare

Gli SNPs sono delle variazioni di un singolo nucleotide nella sequenza del DNA genomico. Come tutte le varianti genetiche, possono presentarsi in omozigosi o in eterozigosi. Il nucleotide, ovvero l'allele che viene studiato in associazione ad una patologia è detto di rischio, a meno che questo non abbia un effetto protettivo. La presenza dell'allele di rischio può essere determinante se questo è presente in entrambi i loci, ma in alcuni i casi anche se presente solo in eterozigosi.

Un SNP che risiede in un esone e porta al cambiamento di un aminoacido, si dice nsSNP (SNP non sinonimo). Se invece l'SNP non porta alla variazione del peptide allora è detto sSNP (SNP sinonimo). Gli SNPs sono classificabili anche in base alla loro posizione e relazione con la patologia. Alcuni SNPs risiedono in promotori o regioni 3' o 5' UTR, essi agiscono direttamente incrementando i livelli di alcuni fattori di crescita o citochine. Altri provocano mutazioni missenso. In questi casi rientrano in una categoria di SNPs detti funzionali, caratterizzati quindi da un'azione biologica diretta e tangibile. Quando invece un SNP viene associato ad una o più patologie ma la sua posizione, ad esempio su un introne o intergenica, non fornisce informazioni per una sua azione diretta, allora non è un SNP

funzionale. In questo caso sono state fatte diverse ipotesi. In alcuni casi si tratta di SNP funzionali, ma con una funzione ancora non ben determinata. In altri casi queste varianti geniche sono semplicemente associate ad altri elementi genici.

Uno dei più grandi database che raccoglie quasi tutti gli SNPs conosciuti è dbSNP e si trova sul sito web www.ncbi.nlm.nih.gov/SNP/. Ad ogni SNP viene associato un codice numerico preceduto dalla sigla "rs". Il codice, che rende identificabile il polimorfismo su dbSNP, è ormai entrato nella pratica comune e utilizzato di conseguenza in letteratura.

Ogni SNP ha delle frequenze che possono variare parecchio da popolazione a popolazione. Ogni volta che un polimorfismo viene associato ad una particolare patologia è necessaria una certa cautela nel considerarlo un fattore di rischio al livello globale. Finché un'associazione non è ancora confermata, al livello mondiale, attraverso diversi studi, è sensato considerare questa relazione solo nel contesto geografico in cui la si è studiata. In risposta a questo problema sono nati gli studi GWAS (genome-wide association study), che cercano delle associazioni con gli SNPs attraverso un approccio che prevede la raccolta di una grande quantità di dati, su altrettanti soggetti, coinvolgendo diversi centri nel mondo.

Gli SNPs sono una classe di alterazioni genetiche oggi ampiamente studiate. Essi sono stati associati a innumerevoli patologie, tramite studi di prospettiva o studi caso-controllo. Alcuni esempi sono i polimorfismi rs1143627 sul gene *IL1B* e rs2981579 sul gene *FGFR2* che sono stati associati al carcinoma alla mammella (Peng S., *et al.*, 2011) o l'rs181946 sul gene *EGFR* associato al carcinoma al polmone (Choi JE., *et al.*, 2007). Come già accennato sono tanti gli SNPs associati al CE. Tramite uno studio bioinformatico ne abbiamo selezionato e studiato alcuni, seguendo dei precisi criteri che saranno illustrati in seguito.

2.1 L'SNP rs231775 A/G sul gene *CTLA4*

Uno dei polimorfismi analizzati nella presente ricerca è l'rs231775 A/G sul gene *CTLA-4* (cytotoxic T-lymphocyte antigen 4). Questo SNP è un polimorfismo funzionale localizzato sul primo esone del gene *CTLA-4*, che porta alla sostituzione aminoacidica Treonina>Alanina sul residuo 17 del peptide leader. L'rs231775 è stato associato ad alcune patologie autoimmunitarie e a diverse neoplasie, come il carcinoma al polmone, all'esofago, alla mammella e alla cervice uterina, oltre che al CE (Sun T., *et al.*, 2009; Hu L., *et al.*, 2010; Gu X., *et al.*, 2010). Sembra in particolare che l'allele A incrementi il rischio di neoplasie, mentre sia protettivo per il rischio di malattie autoimmunitarie. *CTLA-4* è un gene di circa 6 kb che codifica per una proteina di circa 25 kDa. CTLA-4 è un recettore omologo al CD28 presente sulla superficie dei linfociti T, coinvolto nella regolazione negativa della proliferazione

e dell'attivazione. CTLA-4 interagisce con il recettore B7.1 (CD80)/B7.2(CD86) sulla superficie delle cellule presentanti l'antigene (APCs), inducendo l'arresto del ciclo cellulare e l'inibizione della produzione di citochine. CTLA-4 inoltre controlla indirettamente le cellule T effettrici grazie alla sua espressione costitutiva sui linfociti T regolatori. In generale quindi CTLA-4 media la risposta immunitaria T mediata. Sono state, di conseguenza, fatte delle ipotesi che vedono questo recettore coinvolto nella riduzione della risposta immunitaria ai tumori e nell'immunotolleranza. Si suppone infatti che CTLA-4 possa elevare la soglia di attivazione dei linfociti T, attenuandone la risposta antitumorale e aumentando la suscettibilità al cancro (Sun T., *et al.*, 2009). Questi meccanismi non sono ancora stati del tutto chiariti e richiedono ulteriori attenzioni.

La sostituzione Treonina>Alanina causata dal polimorfismo rs231775 A/G porta ad un incremento dell'affinità di CTLA-4 con B7.1. La conseguenza di quest'alterazione è un rafforzamento dell'abilità di inibire la proliferazione e l'attivazione dei linfociti T (Sun T., *et al.*, 2009). Ulteriori esperimenti hanno confermato che i PBMCs (Peripheral Blood Mononuclear Cells) omozigoti per l'allele A, mostrano un decremento dell'attivazione e della proliferazione delle cellule T in confronto ai possessori dell'allele G (Sun T., *et al.*, 2008).

L'ipotesi è quindi che i gli omozigoti per l'allele A dell'rs231775 A/G abbiano un maggiore rischio di neoplasie per l'effetto inibitorio sulla risposta immunitaria ai tumori di questa variante. Nel tessuto epatico inoltre questo effetto si combina con il già alto livello di immunotolleranza locale, dovuto alla produzione costitutiva di citochine antinfiammatorie (Tiegs G., *et al.*, 2010). Gli omozigoti per l'allele G invece sono più a rischio di malattie autoimmunitarie a causa dell'aumentata immunosorveglianza, causata da una riduzione dell'inibizione della proliferazione delle cellule T.

Il polimorfismo in questione è stato studiato prevalentemente nelle regioni asiatiche. Ad oggi mancano dei dati riguardanti le popolazioni europee.

2.2 L'rs2304052 T/C sul gene SPARC

Il polimorfismo rs2304052 T/C è un sSNP sul gene osteonectina/SPARC (secreted protein, acidic, cysteine-rich). L'SNP porta ad una variazione nucleotidica che non corrisponde ad una variazione aminoacidica. La sua funzione biologica è ancora sconosciuta. Questo SNP è stato associato al CE in una popolazione dell'Italia del nord (Segat., *et al.*, 2009) e non è il primo SNP associato al cancro trovato sul gene SPARC. Un altro SNP è, ad esempio, quello presente in posizione 3'UTR, studiato nel 2007, e che porta all'instabilità dell'mRNA di SPARC (Aouacheria A., *et al.*, 2007).

SPARC è un gene di 25.9 kb localizzato sul cromosoma 5, che codifica per una proteina di 32 kDa, costituita da tre domini.

SPARC è una proteina espressa nella matrice extracellulare che ha diversi ruoli fisiologici nella differenziazione cellulare, nella rigenerazione delle ferite, nella regolazione della matrice extracellulare, dell'espressione di fattori di crescita e di alcuni meccanismi di segnalazione intracellulare (Chlenski A., *et al.*, 2010). In particolare SPARC interagisce con diversi componenti della matrice extracellulare. Quando interagisce con il collagene, ad esempio, SPARC porta alla formazione di fibrille in maniera spontanea *in vitro*, agendo così da chaperon molecolare. SPARC inibisce l'angiogenesi, interagendo indirettamente con il bFGF e direttamente sul VEGF. Probabilmente SPARC interagisce con l'integrina β_1 portando alla riduzione della formazione di fibrille di actina *in vitro*. Si documentano anche interazioni tra SPARC, la Stabilina-1 e VCAM-1. L'attivazione di vie di segnalazione coinvolte in parecchie neoplasie, come quelle PI3K/akt, MAPK e Wnt è un'altra proprietà di SPARC (Chlenski A., *et al.*, 2010). Uno degli effetti dell'assenza di SPARC è il sensibile decremento nella massa ossea in topi SPARC^{-/-}, questo dimostra il suo ruolo fondamentale nello sviluppo (Bradshaw AD., *et al.*, 2001).

Nel cancro SPARC sembra avere un ruolo nella transizione epitelio-mesenchimale, contribuendo ad esempio, alla trasformazione dei cloni di melanociti a melanomi (Fukuzanga-Kalabis M., *et al.*, 2007). L'azione di SPARC è più in generale legata alle fasi più avanzate di una neoplasia, è associata infatti, soprattutto alla progressione tumorale e al processo di metastatizzazione (Chlenski A., *et al.*, 2010). Nel CE in particolare c'è un incremento dell'espressione di SPARC che correla con l'angiogenesi tumorale (Lau CP., *et al.*, 2006).

Date le diverse funzioni che SPARC può avere nel cancro, chiarire il coinvolgimento del polimorfismo rs2304052 T/C nel CE potrebbe offrire interessanti spunti per future ricerche e diventare un potenziale fattore predittivo. Non ci sono inoltre ancora dati relativi alla popolazione siciliana.

2.3 L'rs4444903 G/A sul gene *EGF*

L'SNP rs4444903 è un polimorfismo in posizione 61 sul 5'UTR del gene *EGF* (epidermal growth factor). L'rs4444903, conosciuto anche come 61*G, è stato associato a parecchie neoplasie, come il glioma, il cancro allo stomaco, al polmone e il CE. L'associazione con il CE è stata ottenuta indipendentemente su un campione nella popolazione nordamericana e su uno nella popolazione francese, rispettivamente mediante uno studio di prospettiva e uno caso-controllo (Abu Dayyeh BK., *et al.*, 2011, Tanabe KK., *et al.*, 2008). Il 61*G è un polimorfismo funzionale, i portatori in omozigosi dell'allele G, hanno livelli sierici di EGF maggiori rispetto ai soggetti A/A (Abu Dayyeh BK., *et al.*, 2011). La variante sembra infatti influenzare l'efficienza di trascrizione di *EGF*.

Il microambiente pro-tumorale che si forma in stati di infiammazione epatica cronica, come la cirrosi, è caratterizzato dall'azione di diverse proteine della matrice extracellulare, citochine e fattori di crescita, uno di questi è proprio il fattore di crescita epidermico (EGF) (Jang JD., *et al.*, 2011). Il gene *EGF* è lungo circa 100,000 bp e codifica per una proteina a basso peso molecolare. EGF agisce legandosi al suo recettore EGFR, che dimerizza, attivando la via delle MAP chinasi. Questa via promuove la proliferazione cellulare. Nel tessuto epatico in particolare, EGF è considerato un potente mitogeno. L'EGF viene prodotto dalle stesse cellule maligne con un meccanismo di feedback autocrino. Esso ha un ruolo anche nella determinazione della morfologia degli epatociti e si ipotizza che la sua espressione sia correlata con le forme criptogenetiche di CE (Borlak J., *et al.*, 2005). Sembra che EGF giochi un ruolo nell'epatite cronica indotta dall'infezione con *Helicobacter hepaticus*. Da alcuni studi di overespressione di EGF in modelli murini si evidenzia l'incremento dell'espressione di parecchi geni, tra cui *eps-15*, un substrato di EGFR e di *c-fos*, un fattore di crescita attivamente coinvolto nella via di segnalazione attivata da EGF (Borlak J., *et al.*, 2005). Anche nel caso dell'rs4444903 G/A non ci sono dati in letteratura su un'eventuale relazione con il CE nella popolazione siciliana. Questo è infatti un altro dei potenziali marcatori inclusi in questa ricerca.

2.4 L'rs430397 A/G sul gene *GRP78*

La variante genetica rs430397 risiede nella porzione 3' terminale del quinto introne del gene *GRP78* (Glucose-Regulated Protein, 78kDa). Questo polimorfismo è stato recentemente associato, tramite studi caso-controllo, nella popolazione cinese, alla cirrosi epatica e al CE (Zhu X., *et al.*, 2009; Zhu X., *et al.*, 2011). Le implicazioni biologiche dell'rs430397 sono ancora sconosciute. In particolare i portatori dell'allele A, anche in eterozigosi hanno un rischio maggiore di CE rispetto agli omozigoti GG. Questo rischio aumenta ulteriormente se si considerano i soggetti HBV positivi e cirrotici. *GRP78* è un gene di circa 6,5 kb che codifica per una proteina di circa 72 kDa che fa parte della famiglia degli chaperons molecolari. Gli chaperons molecolari sono una classe funzionale di famiglie proteiche, la cui funzione predominante è la prevenzione di associazioni non corrette e aggregazione di catene polipeptidiche non ripiegate, sia in condizioni fisiologiche che in condizioni di stress. *GRP78* si attiva in particolare in risposta ad uno stress del reticolo endoplasmatico (RE). Quando il reticolo endoplasmatico eccede la sua capacità di "ripiegare" le catene polipeptidiche nascenti si attiva un complesso proteico detto UPR, costituito da tre componenti: PERK, IRE1 e ATF6. Questi tre enzimi sono normalmente sequestrati da *GRP78*, in caso appunto di stress del reticolo, *GRP78* si dissocia dal complesso UPR, che è così libero di agire. Il risultato finale dell'attivazione di

UPR è quello di impedire ulteriore ingresso di polipeptidi nel RE e di ridurre la sintesi proteica. In caso di stress del RE prolungato, UPR ha anche una funzione proapoptotica, attraverso la via di CHOP/GADD153 (Zhang L-H., *et al.*, 2010). GRP78 si trova oltre che nel RE, anche nel citosol, prevalentemente nella regione del RE perinucleare. Esso partecipa in questa localizzazione al corretto ripiegamento dei polipeptidi, facilitando la degradazione dei polipeptidi abortivi o non correttamente ripiegati. GRP78 è anche capace di sequestrare il Ca^{2+} intracellulare, infatti controlla l'omeostasi del Ca^{2+} nel RE. Si pensa che questa sua attività sostenga la sopravvivenza cellulare (Zhang L-H., *et al.*, 2010).

Nel cancro la funzione di GRP78 non è ancora chiara. GRP78 può trovarsi anche sulla superficie delle cellule tumorali, può essere addirittura tagliata da alcune metalloproteasi e rilasciata all'esterno della cellula, può essere infatti rilevata nel siero dei pazienti affetti da CE (Zhang L-H., *et al.*, 2010). Dalla superficie cellulare GRP78 riconosce una serie di ligandi, come l' $\alpha 2$ -macroglobulina o la T-caderina, attivando vie di segnalazione che promuovono la sopravvivenza e la proliferazione cellulare. Ma GRP78 sulla superficie cellulare sembra avere anche un'attività proapoptotica. Esso infatti interagisce con l'inibitore dell'angiogenesi plasminogeno Kringle5, inducendo effetti apoptotici (Zhang L-H., *et al.*, 2010). Ancora il GRP78 di superficie può complessarsi con Cripto, un importante fattore coinvolto nell'embriogenesi e nella progressione tumorale.

Il polimorfismo rs430397 A/G sul gene *GRP78*, potrebbe essere un potenziale marcatore di CE nella popolazione siciliana, per questo e per cercare di chiarire il ruolo di GRP78 nel cancro è stato incluso nella presente ricerca.

2.5 L'rs738409 G/C sul gene *PNPLA3*

L'rs738409 è un nsSNP sul gene *PNPLA3*. Questo polimorfismo porta alla variazione aminoacidica isoleucina>metionina in posizione 148 della proteina PNPLA3 (patatin-like phospholipase domain containing 3), detta anche adiponutrina. L'SNP è stato associato a diverse patologie che colpiscono il fegato come la steatoepatite, la cirrosi alcolica, la NAFLD e, anche se in maniera ancora non del tutto chiara al CE (Nischalke HD., *et al.*, 2011). Oltre ad essere stata associata al rischio di insorgenza di queste patologie PNPLA3 è correlata anche ad una peggiore prognosi in alcuni casi di CE (Valenti L., *et al.*, 2013).

PNPLA3 è un gene di circa 40 kb il cui prodotto è una proteina transmembrana con un peso approssimativo di 53 kDa, espressa nell'uomo soprattutto nel fegato e nel tessuto adiposo. PNPLA3 è costituita da un dominio N-terminale patatino(iPLA2)-simile con attività di lipasi trigliceridica e acil-coA transacilasi. L'adiponutrina sembra avere una funzione nell'omeostasi del metabolismo

lipidico, anche se la sua attività fisiologica non è ancora chiara. Di conseguenza è ancora oggetto di discussione il ruolo che gioca nelle patologie. Una delle principali diatribe riguarda l'attività enzimatica dell'adiponutrina, alcuni studi dimostrano un'attività lipasica e altri una transacilasi. Un'altra ipotesi è che svolga entrambe le funzioni in base alla concentrazione del substrato (Valenti L., *et al.*, 2013). Anche sul ruolo del polimorfismo rs738409 sono state fatte diverse ipotesi. Alcuni studi dimostrano una perdita di funzione, mentre altri un guadagno di funzione, a favore di un'attività esclusivamente epatica di sintesi trigliceridica (Valenti L., *et al.*, 2013). Un'altra ipotesi è che PNPLA3 idrolizzi i trigliceridi contenuti nelle "goccioline" lipidiche rilasciando acidi grassi, che diventano così disponibili per la sintesi di trigliceridi. La variante di rischio di PNLA3 quindi potrebbe interferire proprio con quest'attività, riducendo il turnover intracellulare lipidico e l'esportazione dagli epatociti (Valenti L., *et al.*, 2013).

L'associazione del polimorfismo in questione con la steatosi è stata ampiamente confermata, ma, i possessori della variante genetica hanno anche un rischio maggiore di NAFLD, cirrosi da infezione cronica da HCV e CE (Nischalke HD., *et al.*, 2011). Si ipotizza infatti che l'rs738409 su PNPLA3 abbia anche un'attività proinfiammatoria e cancerogenetica diretta, tramite la promozione dell'infiammazione epatica, l'alterazione del rilascio di alcune adipochine, lo stress ossidativo dovuto ad un eccesso di perossidazione lipidica e l'incremento della lipogenesi che potrebbe supportare una rapida crescita cellulare (Valenti L., *et al.*, 2013).

Il polimorfismo in questione è stato studiato ampiamente nella popolazione europea, ma mancano ancora dei dati riguardanti la popolazione siciliana e sono carenti i dati di associazione con il CE. L'rs738409 è stato infatti incluso nel nostro studio.

3. Epigenetica e metiloma nel carcinoma epatocellulare

La metilazione, ovvero l'aggiunta di un gruppo metile, è l'unica modifica naturale del DNA che caratterizza i mammiferi e la metilazione sul sito citosina-fosfato-guanina (CpG) è la meglio caratterizzata. Un dinucleotide CpG si trova nel DNA genomico umano con una frequenza di uno ogni 80 nucleotidi. Esistono delle cosiddette isole CpG dove questi dinucleotidi si concentrano (lunghe dalle 200bp a diverse kb) e hanno una frequenza dell'1-2% nel genoma umano. Spesso queste isole si concentrano nei pressi dei promotori di alcuni geni, si stima che almeno il 50-60% dei geni abbiano delle isole CpG associate al proprio promotore (Sceusi EL., *et al.*, 2011). La metilazione eccessiva o ipermetilazione su sequenze CpG che si trovano su promotori può impedire o rendere più difficile il legame di attivatori o repressori della trascrizione, alterando così l'espressione di quel

gene. Al contrario invece una riduzione dello stato di metilazione naturale su un promotore porta alla sua ipometilazione, causando altrettante anomalie nell'espressione di quel gene. La metilazione al 5' di una citosina è una reazione catalizzata dall'enzima DNA metiltrasferasi (DNMT), la reazione opposta è catalizzata invece dalle demetilasi, come la demetilasi istonica.

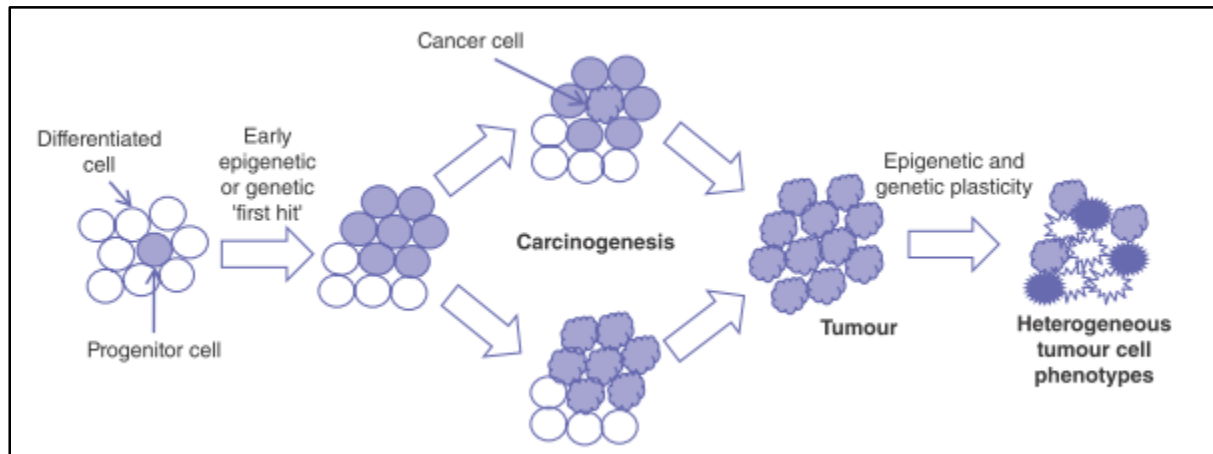


Figura 4: Modello di epatocarcinogenesi mediata da precoci alterazioni epigenetiche (Sceusi EL., et al., 2011).

Le modifiche epigenetiche come la metilazione potrebbero essere uno dei due avvenimenti postulati dall'ipotesi del "two hits" di Knudson. I pattern epigenetici sono stabiliti durante lo sviluppo e sono differenti nei diversi tessuti e tipi cellulari. Le alterazioni epigenetiche potrebbero contribuire quindi insieme alle mutazioni genetiche allo sviluppo di una neoplasia, come schematizzato in figura 4. Un esempio è il silenziamento epigenetico di *MGMT*, un gene deputato al riparo al DNA, riscontrato in neoplasie all'esofago, al rene e alla mammella (Sceusi EL., et al., 2011). In questa direzione infatti si stanno orientando diversi studi sulla caratterizzazione epigenetica delle cellule staminali tumorali.

Si è dimostrato inoltre che a diversi tipi di neoplasie corrispondono peculiari impronte epigenetiche. Le neoplasie gastrointestinali presentano, ad esempio, ipermetilazione su *p16*, *p14*, *MGMT* e *APC*, mentre il CE su *GTP1*, *COX-2*, *RASSF1a* e *TIMP3* (Sceusi EL., et al., 2011).

Tra le alterazioni individuate nella cirrosi epatica ci sono anche quelle epigenetiche. Si ritrovano infatti ipermetilati i promotori di *RASSF1a*, *GSTP1* ed *E-caderina* (Sceusi EL., et al., 2011). Sarebbe di grande utilità clinica un marcatore epigenetico che caratterizzi al livello molecolare i CE più precoci. Le alterazioni nel metiloma possono anche essere indotte da fattori ambientali. La proteina di HBV Hbx, causa l'iperespressione di DNMT, attraverso la sua attività di transattivatore trascrizionale, promuovendo indirettamente l'ipermetilazione genomica (Herceg Z., et al., 2011). Nei CE HCV

positivi, invece, sono più frequenti le alterazioni epigenetiche nei geni *SOCS-1*, *STAT1* e *APC* (Herceg Z., *et al.*, 2011).

I cambiamenti nello stato di metilazione di specifici geni sono rilevabili già in stadi molto precoci, ad esempio nell'epitelio di soggetti affetti da infezioni croniche da HCV e HBV e nel siero e nelle urine di pazienti affetti da CE (Sceusi EL., *et al.*, 2011). I geni più colpiti da metilazione aberrante nel CE sono *RASS1Fa*, *CRABP1*, *GSTP1*, *CHRNA3*, *DOK1*, *SFRP1*, *GADD45*, *p16* e *p15*. Negli stadi più avanzati di CE invece i geni più frequentemente ipermetilati sono *CHFR* e *SYK* (Herceg Z., *et al.*, 2011).

È evidente, dal quadro fin qui disegnato, che la metilazione del DNA nel CE può essere un marcatore di malattia precoce. La proteina TFPI2 (methylated tissue factor pathway inhibitor 2), ad esempio, è rilevabile su pazienti allo stadio I-III CRC con una sensibilità del 76-89% e una specificità del 79-93%. Si stanno studiando inoltre dei chip microarray al fine di verificare uno specifico pattern di metilazione per il CE precoce (Sceusi EL., *et al.*, 2011).

Un altro aspetto da non sottovalutare è l'ipometilazione di specifici geni. L'enzima MBD2 (Methyl-CpG Binding Domain Protein 2) è un agente demetilante che è stato già associato alla demetilazione di specifici geni nel cancro alla mammella e alla prostata (Stefanska B., *et al.*, 2011). MBD2 è stato inoltre associato alla demetilazione e iperespressione dei geni prometastatici *PLAU* e *MMP2* e il suo silenziamento in linee cellulari provoca inibizione dell'invasività e capacità metastatica (Stefanska B., *et al.*, 2011). Occorre quindi una certa cautela nel considerare delle opzioni terapeutiche basate su agenti demetilanti. È necessario chiarire ancora diversi aspetti nel metiloma nel cancro e del CE e fare luce sui meccanismi che governano l'iper- o ipometilazione di specifici geni per poter sviluppare delle terapie mirate e individuare dei validi marcatori di malattia.

Obiettivi

Ad oggi la prevenzione in campo oncologico ha un'importanza sempre maggiore. Lo spettro di terapie disponibili per una malattia ad uno stadio avanzato è ridotto e le prognosi sono in questi casi sempre peggiori rispetto ad una neoplasia diagnosticata in fase precoce.

Solo per poche neoplasie sono stati individuati dei marcatori di malattia più o meno specifici e la loro solidità è continuamente discussa. Al momento attuale non vi è nessun marcatore biologico dotato di sufficiente specificità e sensibilità per il carcinoma epatocellulare (da qui in poi CE)

La ricerca di base ha un ruolo fondamentale nell'individuazione di nuovi marcatori di malattia precoce. Per questo motivo il biotecnologo interagisce sempre di più con il clinico per scegliere dei percorsi di ricerca multidisciplinari e produttivi. Le varianti genetiche polimorfiche sono alcuni dei possibili candidati per svolgere il ruolo di bio-marcatori. Per ovvi motivi la loro efficacia è tuttavia strettamente correlata alla popolazione in cui le si esamina.

Lo scopo di questo lavoro è stato di chiarire il ruolo di alcune varianti genetiche ed epigenetiche nel CE, nella popolazione siciliana, al fine di individuare quelle capaci di predire la suscettibilità alla neoplasia. Per questo, sono stati raccolti parecchi campioni di DNA provenienti da soggetti affetti da CE e soggetti sani e si è messo in pratica uno studio caso-controllo. Alla luce del quadro ottenuto, si è indagato più a fondo su una delle varianti scoperte come fattore di rischio per il CE e sulla situazione epigenetica del gene che la contiene, per chiarire gli aspetti molecolari della nuova evidenza sperimentale nello sviluppo del CE.

Risultati

1. Selezione dei polimorfismi a singolo nucleotide

Mediante una ricerca bioinformatica, condotta sul sito web <http://www.ncbi.nlm.nih.gov/snp>, sono stati scelti 4 polimorfismi a singolo nucleotide (SNPs): rs2304052 T/C, rs4444903 A/G, rs231775 A/G, rs430397 G/A. Inoltre a seguito di una collaborazione con i laboratori del Rockefeller Building dell'University College di Londra, dove era già studiato, è stato analizzato anche il polimorfismo rs738409 G/C.

2. Raccolta, selezione dei campioni e progettazione dello studio

Parte della presente ricerca è stata condotta mediante uno studio caso-controllo. Per questo sono stati raccolti campioni ematici di 365 soggetti sani (controlli), di 175 soggetti affetti da CE (casi) e di 51 soggetti affetti da cirrosi epatica (casi). Da ciascun campione è stato estratto il DNA, la cui qualità è stata saggiata attraverso NanoDrop™ e separazione su gel di agarosio (fig. 5). Sono stati selezionati solo i campioni con una qualità del DNA sufficiente a portare avanti una reazione di PCR-RFLP.

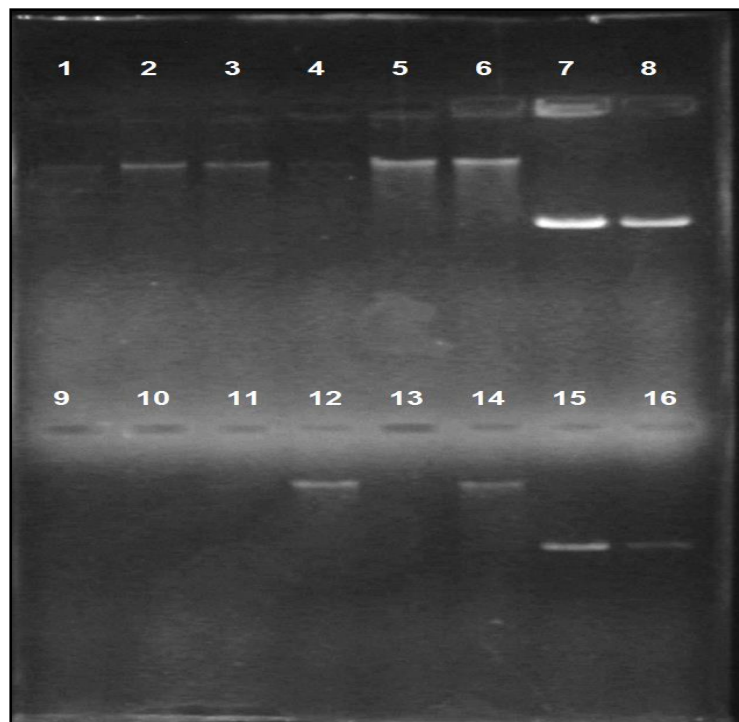


Figura 5: gel di agarosio allo 0,8% colorato con etidio bromuro. Nelle posizioni da 1 a 6 e da 9 a 14 alcuni dei campioni di DNA genomico in analisi. In posizione 7,8,15 e 16 marcatori di quantità con concentrazione rispettivamente di 280,140,70 e 35ng/ul.

Per i controlli sono stati raccolti alcuni dati, quali sesso e età; ci si è accertati inoltre dell'assenza di patologie. Per i casi sono stati raccolti dati riguardanti il sesso, l'età e l'eziologia della patologia. I dati appena menzionati sono riassunti nella tabella 1. La media età dei soggetti sani è di 55,7, con un rapporto uomo:donna di circa 2,27:1. I soggetti affetti da CE hanno una media età alla diagnosi intorno ai 70 anni, con un rapporto uomo:donna di 1,4:1 circa. I casi di cirrosi hanno una media età di quasi 58 anni e il rapporto uomo:donna in questo caso è di 1,5:1.

		eziologia						
		media età	uomini	donne	hcv	hbv	alcool	criptogenetica
casi	controlli	55,7	69,47%	30,53%				
	CE	70,23	57,95%	42,05%	88,53%	2,55%	5,09%	3,82%
	cirrotici	57,65	60%	40%	74,5%	1,96%	15,6%	5,8%

Tabella 1.

3. Genotipizzazione mediante PCR-RFLP

I campioni di DNA genomico inclusi nello studio sono stati genotipizzati per i polimorfismi a singolo nucleotide rs2304052 T/C sul gene *SPARC*, rs4444903 A/G sul gene *EGF*, rs231775 A/G sul gene *CTLA4*, rs430397 G/A sul gene *GRP78*. L'analisi è avvenuta mediante PCR-RFLP (restriction fragment lenght polymorphism): dopo controllo dell'avvenuta amplificazione (vedi fig. 6) i campioni sono stati sottoposti alla fase di restrizione enzimatica e successivamente analizzati su gel di agarosio al 2%, che permette di discriminare per ciascun polimorfismo i tre differenti genotipi, come evidenziato dalle figure 7-10.

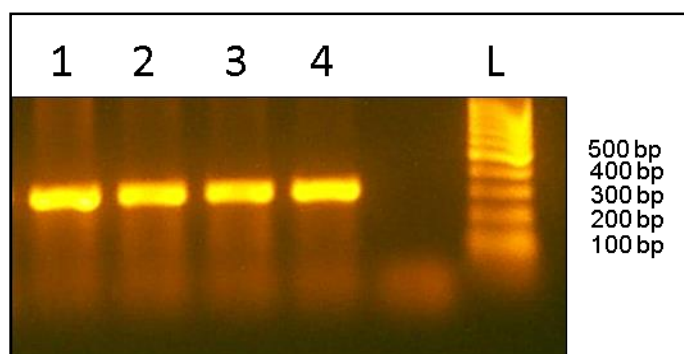


Figura 6: gel di agarosio al 2% colorato con Gel Red®. Prodotti di PCR contenenti l'SNP rs430397 (colonne da 1 a 4). L: DNA ladder da 100bp.

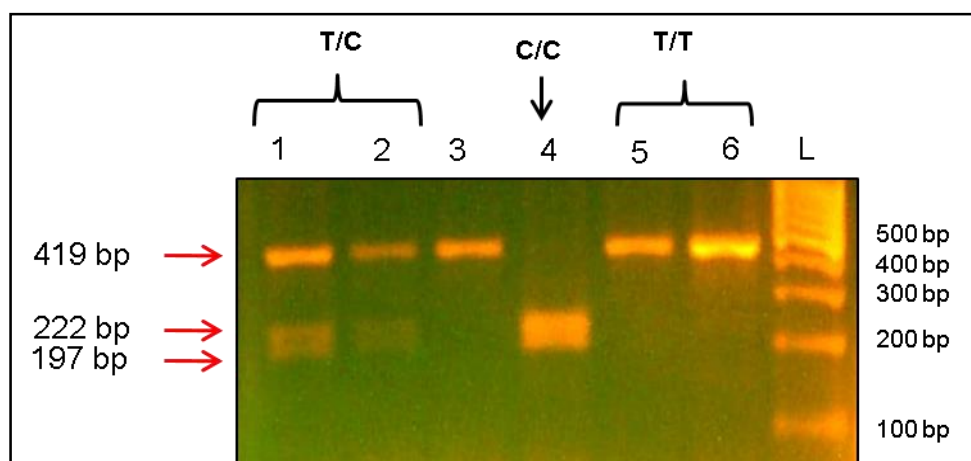


Figura 7: gel di agarosio al 2% colorato con Gel Red®. Separazione dei prodotti di restrizione enzimatica per la discriminazione dei genotipi dell'rs2304052 T/C. Eterozigote T/C (colonne 1 e 2), omozigote C/C (colonna 4), omozigote T/T (colonne 5 e 6). L: DNA ladder da 100bp.

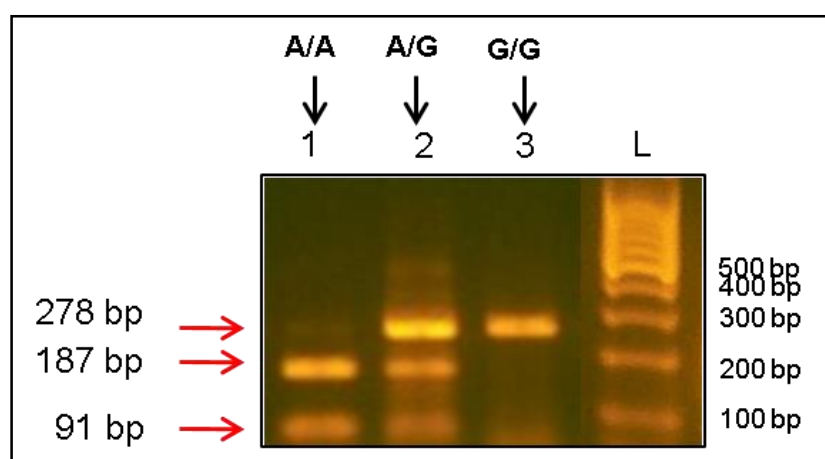


Figura 8: gel di agarosio al 2% colorato con Gel Red®. Separazione dei prodotti di restrizione enzimatica per la discriminazione dei genotipi dell'rs4444903 A/G. Omozigote A/A (colonna 1), eterozigote A/G (colonna 2), omozigote G/G (colonna 3). L: DNA ladder da 100bp.

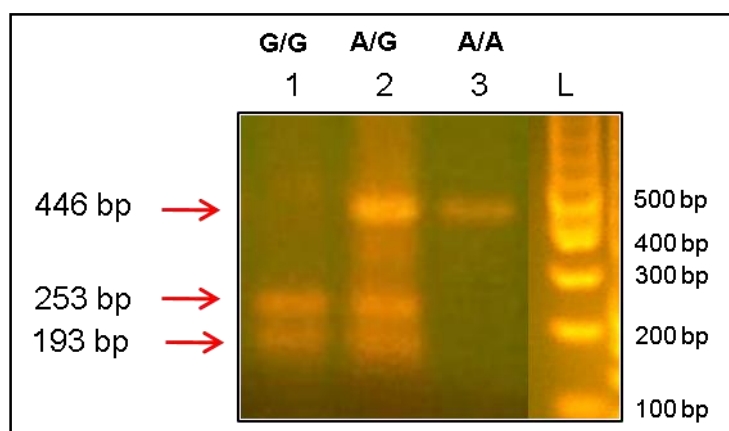


Figura 9: gel di agarosio al 2% colorato con Gel Red®. Separazione dei prodotti di restrizione enzimatica per la discriminazione dei genotipi dell'rs231775 A/G. Omozigote G/G (colonna 1), eterozigote A/G (colonna 2), omozigote A/A (colonna 3). L: DNA ladder da 100bp.

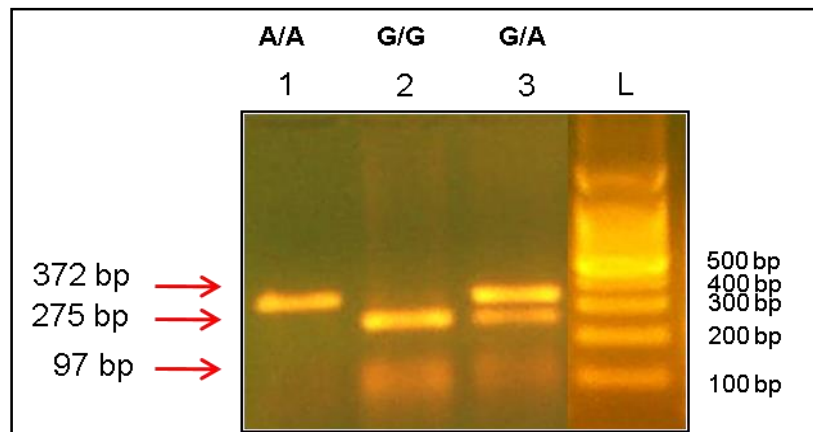


Figura 10: gel di agarosio al 2% colorato con Gel Red®. Separazione dei prodotti di restrizione enzimatica per la discriminazione dei genotipi dell'rs430397 G/A. Omozigote A/A (colonna 1), omozigote G/G (colonna 2), eterozigote G/A (colonna 3), L: DNA ladder da 100bp.

4. Studio caso-controllo: analisi dei dati

I dati raccolti mediante PCR-RFLP, assieme a quelli riguardanti il sesso, l'età e l'eziologia della patologia (già mostrati in tab. 1) sono stati inclusi in un database che è stato poi sottoposto ad analisi statistica. Le frequenze genotipiche degli SNPs analizzati sono in equilibrio di Hardy-Weinberg. Per tutti gli SNPs sono stati calcolati i valori di significatività con il modello classico e sono stati applicati anche i modelli dominante e recessivo. Nel caso di una particolare significatività sono stati calcolati anche gli odds ratios ad un intervallo di confidenza del 95%

Per il polimorfismo rs2304052 T/C sul gene *SPARC*, sono stati analizzati 170 controlli e 75 casi di CE. La frequenza dell'allele minore (minor allele frequency, MAF) è risultata dell'8,8%, bassa rispetto a al 18,8% che troviamo sul database dbSNP. Come è evidente dalla tabella 2 le frequenze genotipiche e alleliche sono molto simili tra casi e controlli. Dall'analisi statistica infatti non è stata evidenziata alcuna significatività, neanche utilizzando i modelli dominante e recessivo. Dopo questa analisi preliminare, che include solo alcuni tra i campioni raccolti abbiamo concluso che non vi è relazione tra il CE e l'rs2304052 T/C sul gene *SPARC* e che i portatori dell'allele C non sono a rischio di CE nella popolazione siciliana.

SNPrs2304052	Controlli (%) n=170	Casi di CE (%) n=75	p
T/T	142 (83,5%)	62 (82,7%)	ns*
T/C	26 (15,3%)	12 (16%)	
C/C	2 (1,2%)	1 (1,3%)	
T	310 (91,2%)	136 (90,7%)	ns
C	30 (8,8%)	14 (9,3%)	
<i>Modello dominante</i>			
T/T	142 (83,5%)	62 (82,7%)	ns
C/C+T/C	28 (16,5%)	13 (17,3%)	
<i>Modello recessivo</i>			
T/T+T/C	168 (98,8%)	74 (98,7%)	ns
C/C	2 (1,2%)	1 (1,3%)	

Tabella 2: analisi statistica sull'SNP rs2304052

*(dato non significativo)

Sono stati genotipizzati tutti i soggetti per l'SNP rs231775 A/G sul gene *CTLA4*. La MAF è risultata essere 30,2%, anche in questo caso più bassa rispetto a quella del dbSNP, che è 45,2%. Nonostante le frequenze genotipiche siano leggermente diverse tra casi e controlli (tabella 3), non sono stati ottenuti dei valori di significatività di interesse, dall'analisi statistica. Anche questo SNP quindi non sembra essere associato al CE nella popolazione siciliana.

SNP rs231775	Controlli (%) n=354	Casi di CE (%) n=163	<i>p</i>
A/A	160 (45,2%)	83 (50,9%)	ns
A/G	174 (49,1%)	70 (42,9%)	
G/G	20(5,7%)	10 (6,2%)	
A	494 (69,8%)	336 (74,7%)	ns
G	214 (30,2%)	80 (25,3%)	
<i>Moodello dominante</i>			
A/A+A/G	334 (94,4%)	153 (93,9%)	ns
G/G	20 (5,6%)	10 (6,1%)	
<i>Modello recessivo</i>			
A/G+GG	194 (54,8%)	80 (49,1%)	ns
A/A	160(45,2%)	83 (50,9%)	

Tabella 3: analisi statistica sull'SNP rs231775

Un altro dei polimorfismi analizzati è l'rs4444903 A/G sul gene *EGF*. La genotipizzazione è stata portata avanti con successo in questo caso per 365 controlli e 154 casi (tabella 4). Le discrepanze nei numeri di casi e controlli analizzati riportati in tabella, rispetto a quelli analizzati per gli altri SNPs, sono da attribuirsi alla degradazione di alcuni campioni avvenuta lungo il corso dell'analisi. In questo caso la MAF ottenuta (50,1%) è più alta rispetto a quella rilevata sul database dbSNP, del 41,8%. Anche per questo SNP si possono notare delle differenze nella distribuzione genotipica su casi e controlli. Dopo il calcolo del valore p con il modello classico, dominante e recessivo, si è ottenuto un dato di non significatività. L'rs4444903 non rappresenta quindi un fattore di rischio per il CE nella popolazione siciliana.

SNP rs4444903	Controlli (%) N=365	Casi di CE (%) n=154	P
A/A	78 (21,4%)	24 (15,6%)	ns
A/G	208 (57%)	101 (65,6%)	
G/G	79 (21,6%)	29 (18,8%)	
A	364 (49,9%)	149 (48,4%)	ns
G	366 (50,1%)	159 (51,6%)	
<i>Modello dominante</i>			
A/A+A/G	286 (78,4%)	125 (81,2%)	ns
G/G	79 (21,6%)	29 (18,8%)	
<i>Modello recessivo</i>			
A/G+GG	287 (78,6%)	130 (84,4%)	ns
A/A	78 (21,4%)	24 (15,6%)	

Tabella 4: analisi statistica sull'SNP rs4444903

L'ultimo SNP analizzato mediante PCR-RFLP su casi e controlli è quello sul gene *GRP78*, l'rs430397.

L'allele A è poco comune in eterozigosi nel nostro campione di popolazione, la MAF è del 9,2%, simile a quella indicata dal database di NCI dbSNP (12%). Dopo l'analisi statistica sono stati ottenuti dei valori di significatività genotipica ($p=0,036$) e allelica ($p=0,023$), sia con il modello classico sia con quello dominante ($p=0,14$). I valori di OR sono 1,61 (95%CI 1,08-2,40) per il modello classico e 1,74 (95%CI 1,13-2,68) per il modello dominante. Se si considerano inoltre solo i soggetti HCV positivi i dati di associazione si rafforzano (p genotipica=0,022; p allelica=0,0038) e gli odds ratios con un intervallo di confidenza del 95% si avvicinano a 2.

SNP rs430397	Controlli (%) n=357	Casi di CE (%) n=170	<i>p</i>	OR (95% CI)	<i>P</i> solo casi HCV+	OR (95% CI) solo casi HCV+
G/G	293(82,1%)	123(72,3%)	0,036		0,022	
A/G	62(17,3%)	46(27,1%)				
A/A	2(0,6%)	1(0,6%)				
Alleli	n=714	n=340				
G	648(90,8%)	292(85,9%)	0,023	1,61 (1,08-2,40)	0,0038	1,92 (1,25-2,96)
A	66(9,2%)	48(14,1%)				
<i>Modello dominante</i>						
G/G	293(82,1%)	123(72,3%)	0,014	1,74 (1,13-2,68)	0,0093	1,79 (1,12-2,84)
G/A+A/A	64 (17,9%)	47 (27,7%)				

Tabella 5: analisi statistica sull'SNP rs430397

5. Genotipizzazione KASPar™

I dati precedentemente mostrati sono stati ottenuti, come più sopra descritto, con la tecnica di RFLP, la quale ha costi limitati, ma notevoli limiti di sensibilità. Si è quindi voluto confermare il dato di significatività attraverso una tecnica più costosa ma molto più sensibile e specifica.

La genoteca in nostro possesso è stata analizzata mediante una tecnica che procede per genotipizzazione diretta e si avvale dell'utilizzo di tre oligonucleotidi e un sistema a doppia fluorescenza FRET, la KASPar™ genotyping (vedere materiali e metodi per di dettagli sulla tecnica). In particolare i campioni di DNA sono stati genotipizzati per il polimorfismo rs430397 e per l'SNP rs738409 G/C sul gene *PNPLA3*. Queste analisi sono state effettuate presso i laboratori del Rockefeller Building dell'University College (UCL) di Londra.

Il saggio KASPar™ ha confermato il 95% dei genotipi ottenuti tramite PCR-RFLP. La successiva analisi statistica ha ribadito i dati di significatività ottenuti precedentemente per l'rs430397 (tabella 6), sia nel modello classico sia in quello dominante. Ancora una volta includendo nell'analisi solo i soggetti HCV+ il dato di significatività subisce un incremento.

SNP rs430397	Controlli (%) N=292*	Casi di CE (%) n=175	P	P solo casi HCV+
G/G	246(83,3%)	129(73,7%)	0,035	0,021
A/G	47(16%)	45(25,7%)		
A/A	2(0,7%)	1(0,6%)		
G	539 (91,4%)	303 (86,6%)	0,02	0,0073
A	51 (8,6%)	47 (13,4%)		
<i>Modello dominante</i>				
G/G	246(83,3%)	129(73,7%)	0,012	0,0087
A/G+A/A	49(16,7%)	46(26,3%)		

Tabella 6: analisi statistica sull'SNP rs430397. Dati ottenuti tramite saggio KASPar™.
 *alcuni dei campioni di DNA non sono stati di qualità sufficiente per l'analisi tramite saggio KASPar™.

Per l'SNP rs738409 G/C sul gene *PNPLA3* la MAF ottenuta è del 30%, simile a quella del 28,4% su dbSNP. L'analisi statistica sui dati ottenuti dalla genotipizzazione KASPar™ hanno mostrato una debole associazione con il CE nel nostro campione di popolazione siciliana (tabella 7). La *p value* allelica è infatti poco al di sotto della soglia di non significatività. Una significatività più marcata è stata ottenuta mediante modello dominante ($p=0,0239$) ma non con quello recessivo. Un altro dato di significatività è stato rilevato tra l'SNP rs738409 e i CE HCV positivi, mediante modello dominante ($p=0,0397$).

I 51 casi di cirrosi sono stati genotipizzati tramite saggio KASPar™ per l'SNP rs430397. Dalla conseguente analisi statistica non si evidenzia nessuna significatività. Sono stati ottenuti *dei p value* non significativi sia al livello allelico, sia genotipico, sia con il modello classico, sia con quello dominante. Il dato di non significatività è confermato anche escludendo dall'analisi tutti i soggetti HCV negativi. L'allele A del polimorfismo rs430397 non sembra quindi influenzare l'insorgenza di cirrosi epatica in soggetti appartenenti alla popolazione siciliana.

SNP rs738409	Controlli (%) N=310	Casi di CE (%) n=177	P	P solo casi HCV+
G/G	152 (49%)	77 (43,5%)	ns	ns
G/C	130 (42%)	72 (40,7%)		
C/C	28 (9%)	28 (15,8%)		
G	434 (70%)	226 (63,8%)	0,048	ns
C	186 (30%)	128 (36,2%)		
Modello dominante				
G/G+G/C	282 (91%)	149 (84,2%)	0.0239	0,0397
C/C	28 (9%)	28 (15,8%)		
Modello recessivo				
G/C+C/C	153 (86,4%)	100 (56,5%)	ns	ns
G/G	152 (13,6%)	77 (43,5%)		

Tabella 7: analisi statistica sull'SNP rs738409. Dati ottenuti tramite saggioKASPar™.

SNP rs430397	Controlli (%) N=298	Casi di cirrosi (%) n=51	P	P solo casi HCV+
G/G	248 (83,2%)	42 (82,4%)	ns	ns
A/G	48 (16,1%)	9 (17,6)		
A/A	2 (0,7%)	0 (0%)		
G	544 (91,3%)	93 (91,2%)	ns	ns
A	52 (8,7%)	9 (8,8%)		
Moodello dominante				
G/G	248 (83,2%)	42 (82,4%)	ns	ns
A/G+A/A	50 (16,8%)	9 (17,6)		

Tabella 8: analisi statistica sull'SNP rs430397 in relazione al rischio di cirrosi epatica. Dati ottenuti tramite saggio KASPar™.

Al fine di indagare sulla natura dell'associazione dell'SNP rs430397 e il CE e per chiarire la sua implicazione nella neoplasia, sono stati individuati 26 casi di CE. Sono stati raccolti i rispettivi 26 tessuti provenienti da resecati chirurgici, conservati in blocchetti paraffinati. Da ciascuno di essi è stato estratto RNA e DNA. Una buona informazione addizionale avrebbe potuto darla l'analisi dell'RNA, ma un controllo della sua qualità ha mostrato uno stato di completa degradazione.

Si è pensato così di passare allo studio del ruolo dell'rs430379 nel CE su linee cellulari. Sono state coltivate le linee cellulari derivate da CE umano: HepG2, Hep3B, Huh7 e SK-HEP1. Le cellule sono cresciute in condizioni basali e, una volta raggiunta la crescita in confluenza è stato isolato l'RNA e il DNA. Entrambe le tipologie di acidi nucleici sono risultate di qualità sufficiente per proseguire con gli esperimenti in programma (figura 11).

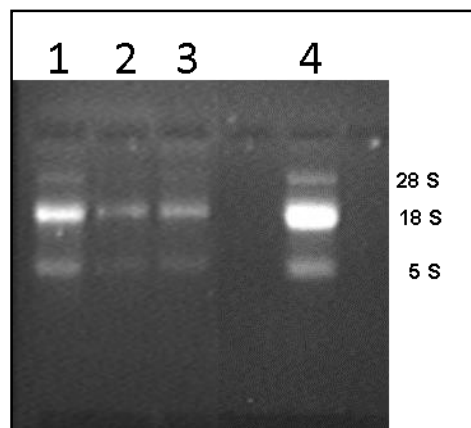


Figura 11: gel di agarosio-formamide all'1,5%, in condizioni denaturanti, colorato con etidio bromuro. Dalla colonna 1 alla 4, nell'ordine, RNA estratto rispettivamente dalle cellule HepG2, Hep3B, Huh7 e SK-HEP1.

6. Valutazione dei livelli di mRNA di GRP78

Al fine di verificare se l'rs430397 sul gene *GRP78* ne possa influenzare l'espressione, l'RNA estratto dalle linee cellulari HepG2, Hep3B, Huh7 e SK-HEP1 è stato retrotrascritto a cDNA. Il cDNA così ottenuto è stato utilizzato come stampo per la reazione di PCR quantitativa. Sono stati disegnati primers specifici per il trascritto di GRP78. L'espressione di GRP78 è stata normalizzata per tre geni housekeeping: la gliceraldeide-3-fosfato deidrogenasi (GAPDH), la succinato deidrogenasi (SDHA) e l'actina beta (ACTB) ed è stata messa in relazione con l'espressione in un tessuto epatico normale (figura 12). Quest'ultimo ha mostrato un'espressione di due volte e mezza rispetto ai geni

housekeeping. Le linee cellulari hanno mostrato un'espressione di GRP78 vicina a quella dei geni housekeeping e di poco inferiore a quella di un tessuto epatico normale. L'espressione di GRP78 nelle cellule HepG2 è invece circa due volte e mezzo quella dei geni housekeeping.

Le linee cellulari sono state genotipizzate per l'rs430397, tramite saggio KASPar™, e sono risultate tutte e quattro omozigoti per l'allele G. Non sembra esserci quindi una relazione tra i livelli di trascritto di GRP78 e l'rs430397.

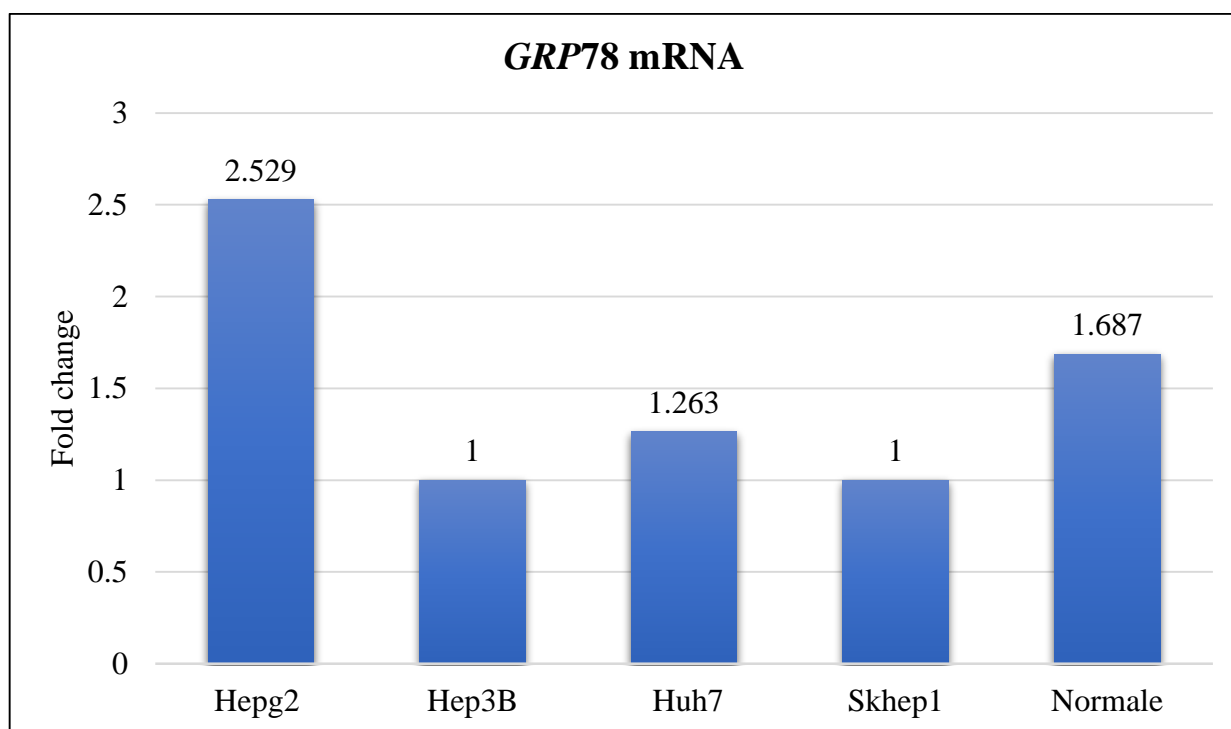


Figura 52. Espressione del trascritto di *GRP78* nelle cellule HepG2, Hep3B, Huh7 e SK-HEP1 e in un tessuto epatico sano. I valori sono normalizzati per la media dell'espressione dei geni housekeeping: gliceraldeide-3-fosfato deidrogenasi (GAPDH), succinato deidrogenasi (SDHA) e actina beta (ACTB).

7. Analisi dello stato di metilazione del promotore di *GRP78*

Le differenze nei livelli di espressione di *GRP78* tra le diverse linee cellulari potrebbero essere dovute ad un diverso stato di metilazione nel promotore di *GRP78*.

Per questo è stato valutato lo stato di metilazione del promotore di *GRP78* delle quattro linee cellulari HepG2, Hep3B, Huh7 e SK-HEP1 e di 7 tessuti affetti da CE, tra quelli precedentemente isolati. Dopo una ricerca in letteratura sullo stato dell'arte delle tecniche per la valutazione dello stato di metilazione di una sequenza di DNA si è scelta la MeDIP (Methylated DNA immunoprecipitation), una tecnica che si basa sull'immunoprecipitazione del DNA metilato. Attraverso una ricerca bioinformatica sono state individuate 6 isole CpG di dimensioni diverse, sul promotore di *GRP78*. In ordine a partire dall'isola più a monte fino a quella che risiede sul sito di inizio della trascrizione le isole contano: 366bp (isola1), 390bp (isola2), 204bp (isola3), 267bp (isola4), 382bp (isola5), 1559bp (isola6). Per motivi economici e tecnici (vedi il capitolo su materiali e metodi) si è deciso di studiare il 50% delle sequenze coperte dalle isole CpG. Per studiare il promotore di *GRP78* si sono scelte: una sequenza CpG sull'isola 1 (1A), una sulla 4 (4B), una sulla 5 (5A) e cinque sulla 6 (6A, 6C, 6D, 6F, 6H), la più grande.

Il grafico 3D in figura 13 mostra tutti i valori di metilazione per ciascun campione e per ogni sequenza CpG analizzata. I valori positivi corrispondono ad ipermetilazione, quelli negativi ad ipometilazione. Per chiarezza nei grafici nelle figure 14 e 15 sono rappresentate rispettivamente le isole ipermetilate e ipometilate. Nel grafico in figura 16 i valori che si riferiscono all'isola CpG 6 sono ricavati dalla media dei valori delle quattro sequenze analizzate per l'isola 6.

I risultati hanno mostrato un generalizzato stato di ipometilazione per i casi di CE, provenienti da resecati chirurgici e uno stato di ipermetilazione per le linee cellulari e nel tessuto epatico normale. Nel grafico in figura 17 (nel quale sono rappresentate le medie dei valori di tutte le sequenze CpG analizzate per ogni campione) questa differenza è evidente. Non si è rilevata alcuna correlazione tra lo stato di metilazione del promotore di *GRP78* sulle cellule HepG2 e l'espressione del trascritto di *GRP78* nelle stesse cellule. Vi è invece una correlazione inversa (coefficiente di correlazione = -0.548) tra lo stato di metilazione e i livelli di trascritto di *GRP78* nelle altre linee cellulari analizzate, le Hep3B, le Huh7 e le SK-HEP.

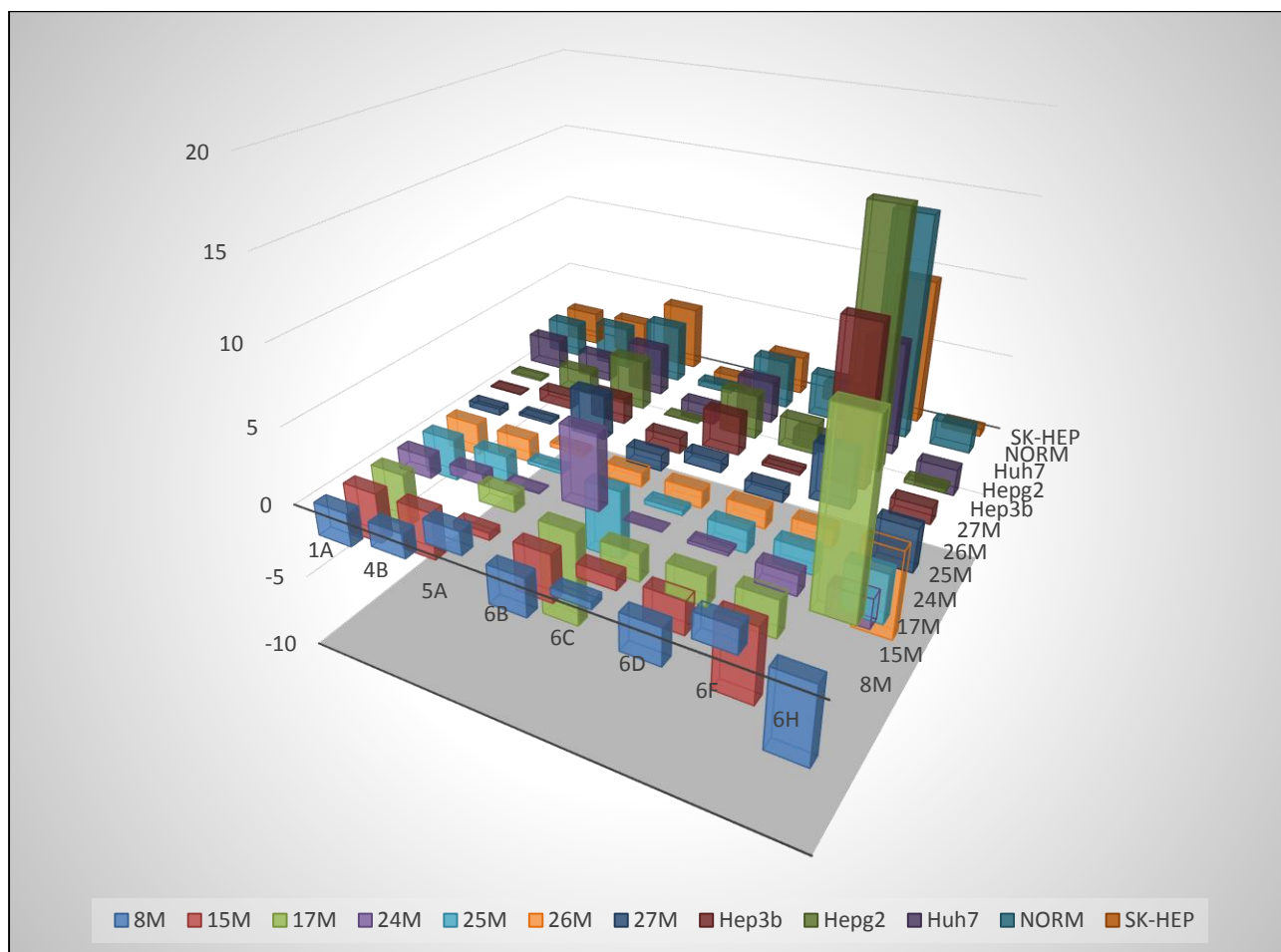


Figura 13. Nei tre assi del grafico 3D sono rappresentati: lo stato di iper/ipo-metilazione (valori positivi o negativi), le sequenze CpG analizzate (1A, 4B, 5A, 6B, 6C, 6D, 6F, 6H, dove il numero si riferisce all'isola CpG e la lettera alla porzione dell'isola analizzata) e i campioni analizzati (da 8M a 27M i tessuti affetti da CE, NORM è un tessuto epatico normale e Hep3B, HepG2, Huh7 e SK-HEP1 sono le linee cellulari analizzate).

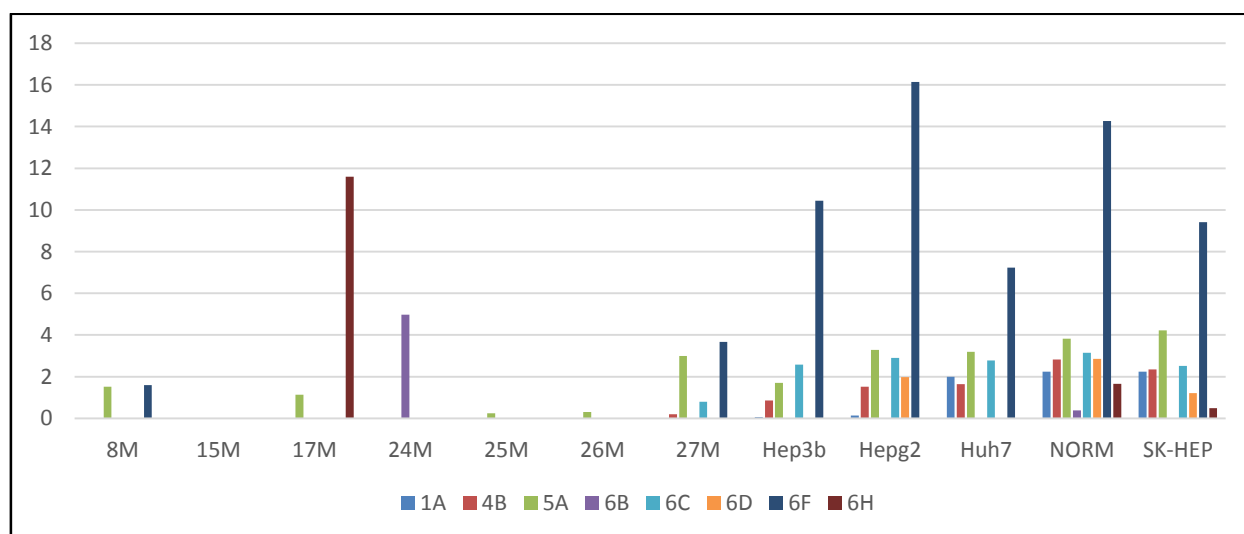


Figura 14: Nel grafico è rappresentato lo stato di ipermetilazione dei campioni analizzati (da 8M a 27M i tessuti affetti da CE, NORM è un tessuto epatico normale e Hep3B, HepG2, Huh7 e SK-HEP1 sono le linee cellulari analizzate) per gli 8 frammenti (da 1A a 6H) che coprono il 50% delle sequenze CpG del promotore di *GRP78*.

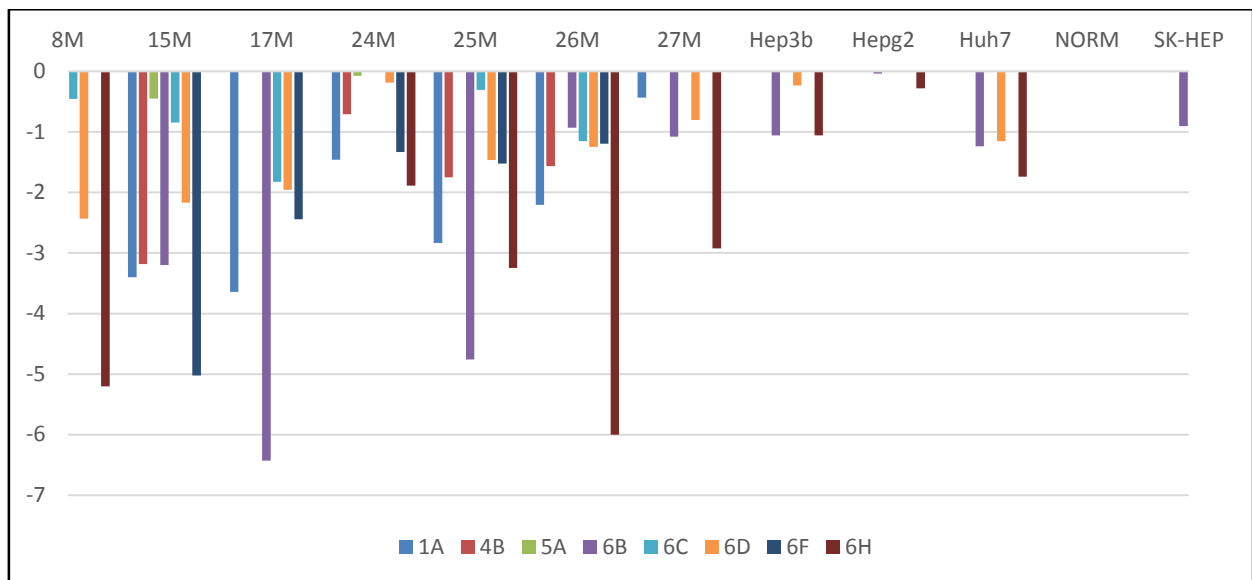


Figura 15: Nel grafico è rappresentato lo stato di ipometilazione dei campioni analizzati (da 8M a 27M i tessuti affetti da CE, NORM è un tessuto epatico normale e Hep3B, HepG2, Huh7 e SK-HEP1 sono le linee cellulari analizzate) per gli 8 frammenti (da 1A a 6H) che coprono il 50% delle sequenze CpG del promotore di *GRP78*.

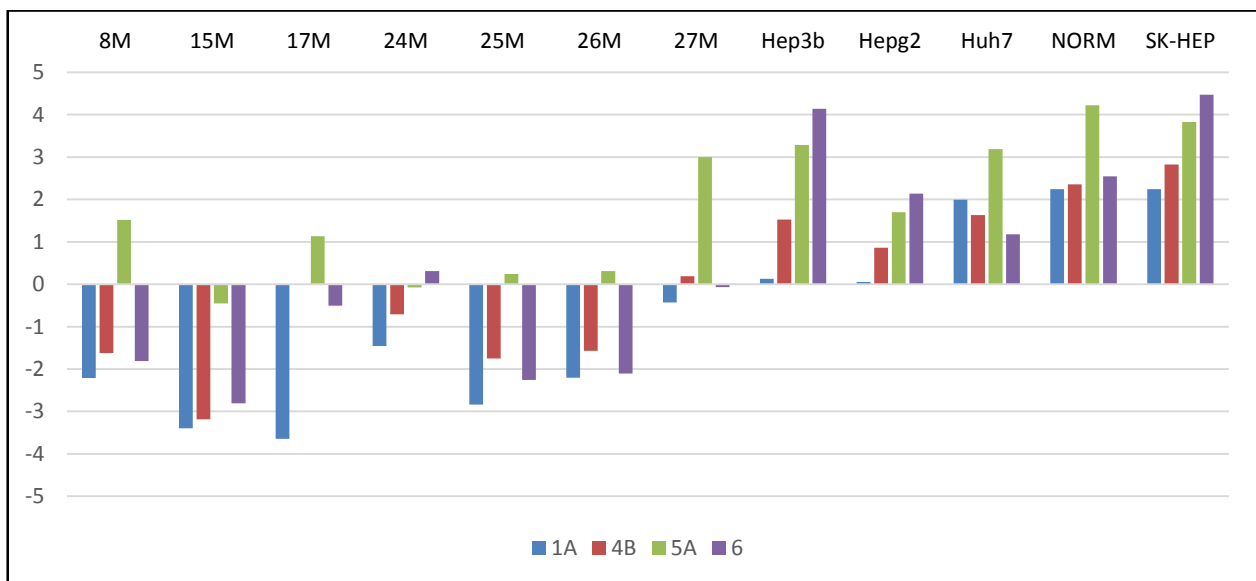


Figura 16. Il grafico mostra lo stato di iper/ipo-metilazione delle sequenze CpG: 1A, 4B, 5A e della media dei valori ottenuti per le sequenze dell'isola CpG 6. I campioni analizzati sono: da 8M a 27M i tessuti affetti da CE, NORM è un tessuto epatico normale e Hep3B, HepG2, Huh7 e SK-HEP1 sono le linee cellulari.

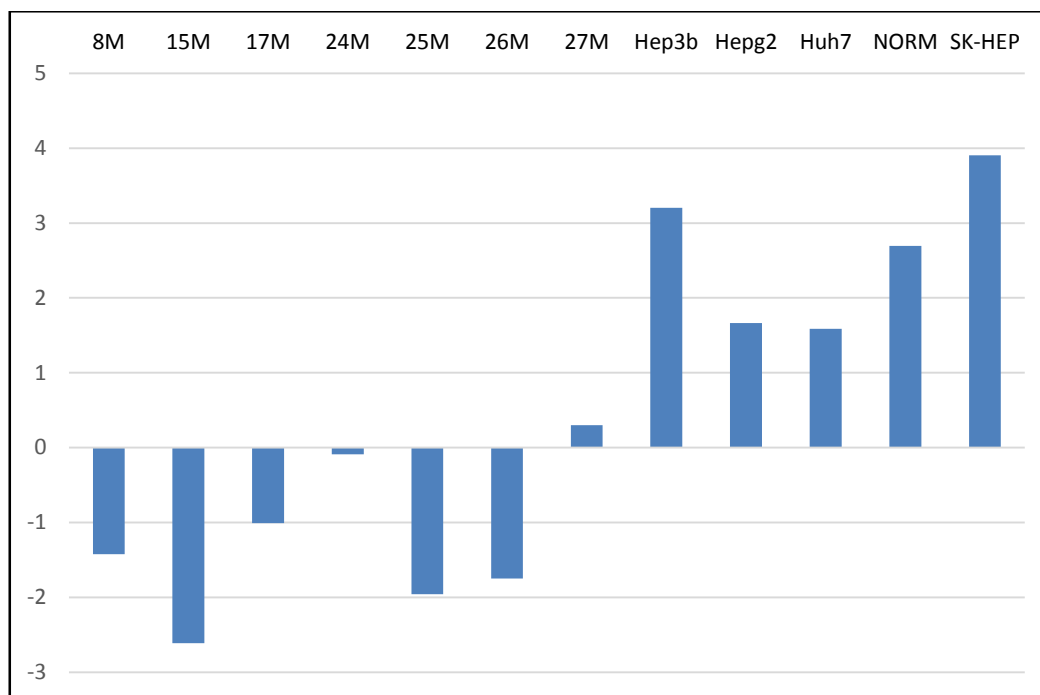


Figura 17. Il grafico mostra lo stato di iper/ipo-metilazione dei campioni: da 8M a 27M i tessuti affetti da CE, NORM è un tessuto epatico normale e Hep3B, HepG2, Huh7 e SK-HEP1 sono le linee cellulari. I valori sono stati ottenuti dalla media dei valori di ciascuna sequenza CpG.

8. Valutazione dell'espressione di GRP78

Al fine di stimare, l'espressione di GRP78, i tessuti sottoposti all'indagine dello stato di metilazione sono stati analizzati tramite immunoistochimica mediante anticorpi specifici per GRP78. L'espressione è stata misurata con dei punteggi, riportati nella tabella 9.

		Casi di CE						
		8M	15M	17M	24M	25M	26M	27M
Punteggi	%cellule positive	3	3	3	3	3	3	3
	Intensità colorazione	2	3	2	2	2	2	2

Tabella 9: in questa tabella sono riportati i punteggi relativi alla percentuale delle cellule positive per GRP78 e all'intensità della colorazione, dopo analisi immunoistochimica mediante anticorpo specifico per GRP78. L'analisi è stata effettuata sulla base di due criteri: percentuale di cellule positive e intensità della colorazione. Per ognuno dei due criteri è stato attribuito un punteggio che va da 1 a 3. Per il primo criterio: 1 equivale dallo 0% al 25% di cellule positive alla colorazione, 2 dal 26% al 50% di cellule positive, 3 a più del 50% di cellule positive. Per il secondo criterio: 0 equivale ad una colorazione negativa, 1 a un'intensità della colorazione debole, 2 a un'intensità moderata, 3 a un'intensità forte. Sommando questi punteggi si ottengono tre classi: un punteggio totale tra 0 e 2 equivale a un campione negativo, un punteggio tra 3 e 4 a un campione debolmente positivo, un punteggio tra 5 e 6 a un campione positivo.

La somma dei punteggi, ha dato il valore massimo (6/6) solo per il campione 15M e un valore, sempre di positività ma leggermente inferiore a quello massimo (5/6), su gli altri 6 campioni. In figura 18 sono riportate alcune delle immagini delle sezioni di tessuto, dopo il trattamento con l'anticorpo specifico per GRP78. La localizzazione di GRP78 è prettamente citoplasmatica.

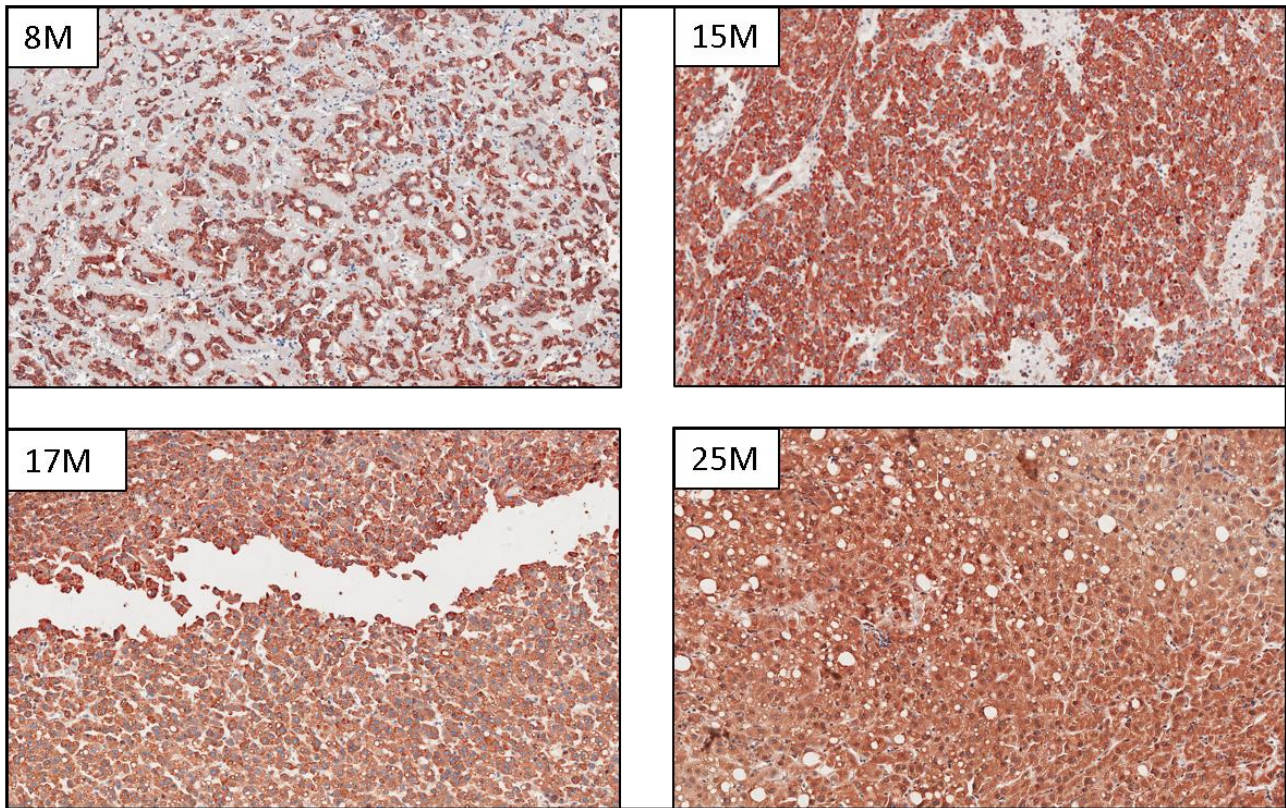


Figura 18: immunoistochimica di quattro dei sette campioni di tessuto di CE sottoposti all'analisi di metilazione. In alto a sinistra sono indicati i codici dei campioni. Le reazioni sono state condotte mediante anticorpo specifico per GRP78.

Pazienti, materiali e metodi

1. Casi di carcinoma, di cirrosi epatica e controlli

Sono stati raccolti 365 campioni ematici di soggetti sani (controlli), 177 di soggetti affetti da CE (casi) e 51 affetti da cirrosi epatica (casi). I controlli sono stati raccolti presso la banca del sangue dell'ospedale ARNAS Civico Di Cristina e Benfratelli. I criteri che hanno portato alla scelta dei soggetti di controllo sono stati: status di donatore ematico sano, età maggiore di 50 anni, residenza in Sicilia, assenza di patologie neoplastiche, valori ematici nella media, assetto lipidico nella norma, HIV, HBV e HCV negatività. Sono stati raccolti dati relativi a sesso ed età per ciascun soggetto.

I casi di CE e di cirrosi epatica sono stati raccolti presso il Dipartimento Biomedico di Medicina Interna e Specialistica (DIBIMIS) del Policlinico Universitario Paolo Giaccone di Palermo. I criteri per la diagnosi di CE e di cirrosi epatica sono quelli utilizzati dal dipartimento nella pratica clinica (Soresi M., *et al.*, 2010). Sono stati raccolti dati virologici, eziologici e dati relativi ad età e sesso per ciascun paziente.

Tutti i soggetti inclusi nello studio hanno firmato un modulo di consenso informato. Il progetto di ricerca è stato approvato dalla commissione etica del Policlinico Universitario Paolo Giaccone di Palermo.

2. Polimorfismi a singolo nucleotide

Lo studio bioinformatico che ha portato alla scelta dei polimorfismi a singolo nucleotide rs2304052 T/C, rs4444903 A/G, rs231775 A/G, rs430397 G/A è stato condotto sui siti <http://www.ncbi.nlm.nih.gov/> e <http://www.ncbi.nlm.nih.gov/snp>, utilizzando le parole chiave: SNP, HCC e case-control study. I criteri che hanno portato a scegliere i quattro SNPs sono stati: associazione con il CE in almeno uno studio con una $p > 0,05$, frequenza dell'allele minore $> 0,5\%$, campione di popolazione in analisi eterogeneo o almeno (casi) non completamente costituito da CE HBV positivi o causati da alcool, SNPs su geni potenzialmente coinvolti nel CE, dati relativi alla variante genetica carenti nelle etnie caucasiche. La MAF (minor allele frequency) globale riportata per ciascun SNP proviene dal sito <http://www.ncbi.nlm.nih.gov/snp>.

3. Estrazione e valutazione della qualità degli acidi nucleici

Ciascun campione ematico è stato subito trattato per l'estrazione del DNA genomico o conservato a -20 °C per un periodo non superiore ad un mese. Prima della conservazione a -20 °C i campioni sono stati centrifugati a 5.000 rpm per 10 minuti e la frazione plasmatica è stata rimpiazzata con un'uguale quantità di soluzione fisiologica di sodio cloruro allo 0,9%.

È stata prelevata una frazione di 300 µl da ciascun campione ematico per l'estrazione del DNA genomico. Sono stati utilizzati i protocolli e i reagenti forniti dai kits commerciali: High Pure PCR TemplatePreparation Kit (Roche©) e Wizard® Genomic DNA Purification Kit (Promega©).

La qualità del DNA genomico è stata valutata tramite comparazione con quattro standard a concentrazione di DNA nota di 280ng/µl, 140ng/µl, 70ng/µl e 35ng/µl. I campioni sono stati separati su gel di agarosio standard allo 0,8%. Sono stati utilizzati 3µl di campione insieme a 0,5 µl di Gel Loading Dye, Blue (6X) (New England Biolabs© inc.). Per valutare l'integrità del DNA genomico inoltre sono stati utilizzati marcatori di peso molecolare: 100bp DNA Ladder (Promega©) e 1kb DNA Ladder (Promega©). Il gel di agarosio è stato sottoposto ad una corrente di 5Volts/cm per 1ora. La qualità del DNA è stata valutata anche tramite lo spettrofotometro NanoDrop® (ThermoScientific™), utilizzando una quantità di 2µl per campione.

L'estrazione di DNA ed RNA dai tessuti paraffinati è stata effettuata grazie al protocollo e ai reagenti forniti dal kit AllPrep® DNA/RNA FFPE (Qiagen©) in condizioni di assenza di RNasi. Superfici e dispositivi sono stati detersi con la soluzione RNase AWAY® (Sigma-Aldrich©).

L'estrazione di DNA ed RNA da colture cellulari e da tessuti epatici sani è stata effettuata simultaneamente, mediante il protocollo del TRIzol (Life Technologies Corporation©). I tessuti epatici sani sono stati gentilmente forniti dal Royal Free Campus dell'University College of London (UCL), dove erano studiati all'interno di un progetto universitario.

La qualità dell'RNA è stata valutata tramite separazione su gel di agarosio-formaldeide all'1,5%, secondo un protocollo già documentato (Sambrook J., *et al.*, 1989). Il gel, immerso in un Buffer di corsa per elettroforesi di RNA, è stato sottoposto ad una corrente di 5Volts/cm per 45 minuti.

4. PCR-RFLP

I dati relativi ai genotipi dei polimorfismi a singolo nucleotide rs2304052 T/C, rs4444903 A/G, rs231775 A/G, rs430397 G/A sono stati raccolti tramite PCR-RFLP (restriction fragment length polymorphism). La PCR-RFLP prevede una fase di amplificazione seguita da una restrizione enzimatica. L'esperimento è stato simulato e progettato elettronicamente grazie allo strumento informatico <http://tools.neb.com/NEBcutter2/>. La fase di PCR è stata condotta utilizzando per ogni SNP gli oligonucleotidi specifici descritti nella tabella 9 e disegnati grazie agli strumenti informatici sui siti web <http://bioinfo.ut.ee/primer3-0.4.0/> e <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>. Per ogni reazione di amplificazione sono state utilizzate le condizioni di reazione specifiche descritte nella tabella 10, ottimizzate per condurre successivamente una reazione di restrizione sul prodotto di PCR. Le reazioni di PCR sono state condotte in un volume di 20 µl, con una concentrazione di DNA di 50 µg, che funge da stampo, e utilizzando la miscela di reagenti con le concentrazioni descritte nella tabella 11. Per verificare i prodotti di PCR, un'aliquota di 5µl del volume di PCR è stata analizzata, insieme a 1µl di Gel Loading Dye, Blue (6X) (New England Biolabs© inc.), su gel di agarosio all'1% in un tampone di TBE 1X. Il gel è stato sottoposto ad una corrente di 5 Volts/cm per 20 minuti. La dimensione dei prodotti è stata accertata grazie all'utilizzo del marcatore di peso molecolare 100bp DNA Ladder(Promega©).

SNP compreso nel prodotto di PCR	Oligonucleotidi
rs2304052	5'-TCAGGAGTGTGTCAGGATGTG-3' 5'-GCTCACATGACTGGCTTCTTC-3'
rs4444903	5'-AGCTTTAAAAAGCAAAGCTGAGT-3' 5'-GTTTAACAGCCCTGCTCTGG-3'
rs231775	5'-CACGGCTTCCTTTCTCGTAA-3' 5'-CAAGCCAGATTGGAGTTTACC-3'
rs430397	5'-GTTCTTGTTTTCTGCTTGCT-3' 5'-GAAAGATGCTGCGATGATGA-3'

Tabella 9: oligonucleotidi utilizzati nella fase di amplificazione dell'RFLP.

SNP compreso nel prodotto di PCR	Cicli	Denaturazione iniziale*	Denaturazione	Annealing	Allungamento	Fase finale*
rs2304052	35	94 °C 6min	94 °C 50 sec	59 °C 50 sec	72 °C 1min	72 °C 6min
rs4444903	35	94,5 °C 4 min	94,5 °C 40 sec	59,5 °C 40 sec	72 °C 50 sec	72 °C 5 min
rs231775	30	95 °C 5 min	94 °C 35 sec	59 °C 35 sec	72 °C 30 sec	72 °C 8 min
rs430397	30	94 °C 5 min	94 °C 35 sec	58 °C 30 sec	72 °C 35 sec	72 °C 6 min

Tabella 10: condizioni di reazione utilizzate per la PCR che precede l'RFLP.

*le fasi di denaturazione iniziale e allungamento finale sono ripetute solo all'inizio e alla fine della reazione

SNP compreso nel prodotto di PCR	Oligonucleotidi	dNTPS	MgCl ₂	PCR buffer (Promega®)	Taq DNA polimerasi (Promega®)
rs2304052	0,5 µM	0,2mM	1,5 mM	1X	0,75 U
rs4444903	0,5 µM	0,2mM	1,5 mM	1X	0,75 U
rs231775	0,4µM	0,2mM	1,8mM	1X	0,5 U
rs430397	0,45 µM	0,2mM	2mM	1X	0,5 U

Tabella 11: condizioni di reazione utilizzate per una reazione di PCR con un volume totale di 20µl e una concentrazione di DNA genomico che funge da stampo di 50 ng/µl.

Le reazioni di RFLP sono state condotte in una miscela di reagenti con un volume finale di 25µl, con le concentrazioni e secondo le condizioni descritte nella tabella 12. Al fine di discriminare tra le diverse forme alleliche, i prodotti di restrizione enzimatica sono stati separati su gel di agarosio al 2% in un tampone di TBE 1X. L'intero volume di reazione è stato analizzato su gel, insieme al Gel Loading Dye, Blue 1X (New England Biolabs® inc.). Il gel di agarosio è stato sottoposto ad una corrente di 4,5 Volts/cm per 45 minuti. La dimensione dei prodotti è stata verificata grazie al marcatore di peso molecolare 100bp DNA Ladder(Promega®). La dimensione dei prodotti ottenuti

mediante restrizione enzimatica, grazie all'utilizzo di enzimi specifici per ogni SNP, ha permesso la discriminazione dei genotipi per ciascun SNP secondo lo schema riportato nella tabella 13.

Miscela di reazione				Condizioni di reazione
SNP compreso nel prodotto di PCR	Prodotto di PCR	Tampone	Enzima	Incubazione
rs2304052	4µl	Neb Buffer 4 (New England Biolabs®) 1X	2 U EcoO109I (New England Biolabs®)	37 °C 16 h
rs4444903	4µl	Neb Buffer 4 (New England Biolabs®) 1X	1 U AluI (New England Biolabs®)	37 °C 8 h
rs231775	4µl	Neb Buffer 3 (New England Biolabs®) 1X	5 U ApeKI (New England Biolabs®)	37 °C 16 h
rs430397	2µl	Buffer K (Promega©) 1X	10 U Hsp92II (Promega©)	75 °C 16 h

Tabella 12: condizioni e concentrazioni dei reagenti dei saggi di RFLP.

SNP		N _{ro} di bande su gel di agarosio al 2%		
		1	2	3
rs2304052 T/C	Dimensione del/i prodotto/i di restrizione	419 bp	222 bp, 197 bp	419 bp, 222 bp e 197 bp
	Genotipo	T/T	C/C	T/C
rs4444903 A/G	Dimensione del/i prodotto/i di restrizione	278 bp	187 bp e 91bp	278 bp, 187 bp e 91 bp.
	Genotipo	G/G	A/A	A/G
rs231775 A/G	Dimensione del/i prodotto/i di restrizione	446 bp	253 bp, 193 bp	446 bp, 253 bp e 193 bp
	Genotipo	A/A	G/G	A/G
rs430397 G/A	Dimensione del/i prodotto/i di restrizione	372 bp	275 bp, 97 bp	372 bp, 275 bp e 97 bp
	Genotipo	A/A	G/G	G/A

Tabella 13: schema secondo il quale sono stati discriminati i differenti genotipi dei quattro SNPs analizzati mediante PCR-RFLP.

5. Colture cellulari

Le linee cellulari derivate da carcinomi epatocellulari umani HepG2, Hep 3b e Huh7 ci sono state cortesemente fornite dall'istituto di Biomedicina e Immunologia Molecolare (IBIM) "Alberto Monroy" del CNR (Centro Nazionale delle Ricerche) di Palermo. La linea cellulare derivata da carcinoma epatocellulare umano SK-HEP 1 ci è stata gentilmente fornita dal Blizzard Institute, Queen Mary University of London di Londra, Regno Unito. Tutte le cellule sono cresciute condizioni standard per i tessuti di mammiferi. È stato utilizzato il terreno DMEM, high glucose (Gibco®), in presenza di 100 U/ml di penicillina e 100 ug/ml di streptomicina per coltivare tutte e 4 le linee cellulari.

6. Genotipizzazione Kaspar™

I saggi KASPar™ (Kompetitive Allele Specific PCR) sono stati condotti seguendo le indicazioni per il protocollo standard sul manuale KASP version 4.0 SNP Genotyping Manual (LGC Genomics©). La miscela di reazione è schematizzata nella tabella 14. Le condizioni di reazione sono quelle standard indicate dal manuale sopracitato. La concentrazione di DNA genomico utilizzata per ogni reazione è stata di 3,3 ng/ul. Il DNA è stato disidratato prima di procedere con il saggio. Gli oligonucleotidi utilizzati sono riportati nella tabella 15. Le reazioni sono state condotte su piastre da 384 pozzetti, avvalendosi del LightCycler® 480 System (Roche Diagnostics©). L'analisi dei dati è stata effettuata grazie al programma di genotipizzazione Kraken™ (LGC Genomics©).

Reagente	µl
2X Reaction Mix	2
Assay	0,055
H2O	2

Tabella 14: miscela di reagenti per un volume di reazione di circa 4µl e per il tipo di reazione "Dry" (il DNA è disidratato).

SNP compreso nel prodotto di PCR	Oligonucleotidi	Marcatura
rs738409	5'-GAAGGTGACCAAGTTCATGCTCCTGGTATGTTCTGCTTCATC-3'	FAM
	5'-GAAGGTCGGAGTCAACGGATTCTTGGTATGTTCTGCTTCATG-3'	HEX
	5'-CGCCTCTGAAGGAAGGAGGGAT -3'	-
rs430397	5'-GAAGGTGACCAAGTTCATGCTAGATAACAGACATCACAGTAACCATG-3'	FAM
	5'- AAGGTCGGAGTCAACGGATTCTTAGATAACAGACATCACAGTAACCATA-3'	HEX
	5'- ATAGTAGACCGGAACAGATCCTAGAAAA-3'	-

Tabella 15: oligonucleotidi utilizzati per il saggio KASPar™. Alcuni di essi hanno una marcatura fluorescente.

7. Real Time PCR quantitativa

L'RNA estratto dalle linee cellulari e dal tessuto epatico sano è stato retrotrascritto seguendo le indicazioni e utilizzando i reagenti forniti dal kit QuantiTect Reverse Transcription Kit® (QIAGEN®).

La reazione di Real Time PCR quantitativa è stata condotta con una metodica basata sul SYBRGreen®, seguendo le istruzioni e utilizzando i reagenti presenti nel kit QuantiFast SYBR® Green PCR Kit(QIAGEN®). Lo strumento utilizzato è lo StepOnePlus™ (Life Technologies Corporation®), compatibile con i kits QIAGEN®.

Il cDNA ottenuto mediante la retrotrascrizione è stato utilizzato come stampo per la reazione di Real Time PCR quantitativa. L'espressione del trascritto di *GRP78* è stata normalizzata per quella dei geni housekeeping: gliceraldeide-3-fosfato deidrogenasi (GAPDH), succinato deidrogenasi (SDHA) e actina beta (ACTB). Sono stati utilizzati gli oligonucleotidi specifici descritti in tabella 16 e disegnati grazie agli strumenti informatici presenti sui siti web <http://bioinfo.ut.ee/primer3-0.4.0/> e <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>. Gli oligonucleotidi sono stati disegnati a cavallo di regioni introniche.

La concentrazione di cDNA nella miscela di reazione è stata scelta grazie ad un esperimento di ottimizzazione dell'efficienza di PCR, nel quale sono state utilizzate, insieme al cDNA puro, diluizioni del cDNA da 1:2 a 1:64. La concentrazione ottimale per un'elevata efficienza di amplificazione è stata stabilita in 1:8.

Ogni reazione è stata condotta in triplice copia su una piastra da 96 pozzetti. L'analisi dei dati è stata effettuata mediante il programma informatico StepOne™ Software v2.2.2 (Life Technologies Corporation©). L'espressione del trascritto di *GRP78* è stata calcolata per mezzo del programma GeneX® (Biorad©).

Gene	Oligonucleotidi per Real Time PCR quantitativa
<i>GRP78</i>	5'-AAATGAAAGAAACCGCTGAGGCT-3' 5'-TGCTTGGCGTTGGGCATC -3'
<i>GAPDH</i>	5'-TGGGAACAAGAGGGCATCTG -3' 5'-CCACCACTGCATCAAATTCATG -3'
<i>ACTB</i>	5'-CACCATGTACCCTGGCATT -3' 5'-CCGATCCACACGGAGTA -3'
<i>SDHA</i>	5'-AGAAGCCCTTTGAGGAGCA -3' 5'-CGATCACGGGTCTATATTCCAGA -3'

Tabella 16: oligonucleotidi utilizzati per la reazione di Real Time PCR quantitativa.

8. MeDip-qPCR (Methylated DNA immunoprecipitation-qPCR)

La MeDip è una tecnica basata sull'immunoprecipitazione del DNA metilato.

Il protocollo che è stato seguito per arricchire il DNA di metilcitosine è stato modificato rispetto a quanto descritto precedentemente (Weber M., *et al.*, 2005). In particolare:

1. Sonicazione del DNA genomico fino a raggiungere una dimensione dei frammenti di circa 150bp. È stato utilizzato lo strumento Bioruptor sonicator (Diagenode, Denville, NJ, USA).
2. Controllo della dimensione del DNA attraverso separazione su gel di agarosio al 3%, secondo quanto descritto nel paragrafo 3.
3. Purificazione del DNA attraverso il PCR purification kit (Qiagen©). Eluizione in 30µl di Buffer EB (provisto dal kit).
4. Diluizione di 500 ng di campione fino ad un volume di 31 µl.

5. Denaturazione a 100 °C per 10 minuti e immediato raffreddamento in ghiaccio per minuti.
6. Aggiunta di 31 µl di tampone IP 2X, 1 µl di anticorpo 5-Methylcytosine (5-mC) Monoclonal Antibody (Cambridge Bioscience Ltd.) e incubazione per 2 ore a 4°C con lieve rotazione.
7. Lavaggio di 4 µl di Dynabeads® M-280 Sheep Anti-Mouse IgG (Life Technologies Corporation©) con 1 ml del tampone IP 1X e recupero delle Dynabeads® mediante dispositivo magnetico.
8. Aggiunta delle Dynabeads® ai campioni e incubazione per 2 ore a 4°C con lieve rotazione.
9. Recupero delle Dynabeads® grazie ad un dispositivo magnetico. Una volta eliminato il supernatante, lavaggio con 500 µl del tampone IP 1X. Ripetere il lavaggio tre volte.
10. Risospensione delle Dynabeads® in 50 µl di tampone di digestione per proteinase K, aggiunta di 1 µl di Proteinase K (recombinant), PCR grade (Thermo Fisher ScientificInc) 10 mg/ml e incubazione per 2 ore a 55°C con rotazione.
11. Raccolta delle Dynabeads® con un dispositivo magnetico. Trasferimento del supernatante in 700 µl del binding buffer del DNA Clean&Concentrator™-5 (ZymoResearch). A questo punto sono state seguite le indicazioni presenti nel kit e i campioni sono stati eluiti in 10 µl of buffer EB.

Una volta preparate le frazioni arricchite di DNA metilato, un'aliquota di queste, insieme ad un'aliquota delle frazioni non arricchite (input) sono state sottoposte a Real Time PCR. Per ogni sequenza CpG sono stati utilizzati degli oligonucleotidi specifici, disegnati grazie agli strumenti informatici presenti sui siti web <http://bioinfo.ut.ee/primer3-0.4.0/> e <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>. Sono stati utilizzati dei controlli positivi e negativi per verificare l'arricchimento delle frazioni dopo la procedura prevista dalla MeDip. Tutti gli oligonucleotidi utilizzati sono riportati nella tabella 17. Ogni reazione di Real Time PCR è stata condotta, in triplice copia, in piastre da 96 pozzetti, su dispositivo 7500 Real-Time PCR System (Life Technologies Corporation©), seguendo le istruzioni e utilizzando i reagenti presenti nel kit Power SYBR® Master Mix (Life Technologies Corporation©).

Sequenza CpG	Oligonucleotidi per Real Time PCR quantitativa dopo protocollo MeDip
1A	5'-CAAGCAATCCTCCTGCCTCAGCCTC -3' 5'-GGGGTTACAGGTGTGAGCCACCG -3'
4B	5'- AGTCCCAGCTACTGCTACTCAGAGG -3' 5'- GTCTCGTTCTGTGGCCAGGCT -3'
5A	5'- GTTTCGAACTCCCAACCTCAAGCGA -3' 5'- GGTCTTGATCTCCTGACCCCGTGAT -3'
6B	5'- TGGCGGCATCGCCAATCAGAC -3' 5'- GGTTCGCGATGCAGGGGGAC -3'
6C	5'- GTAGGTGGTCCCCAGGTCGATGC -3' 5'- GACTGCCTGCTGCTGCCCAAC -3'
6D	5'- GTAGCACAGGAGCACAGCGCAAT -3' 5'- CCGGCCAAGACAGCACAGACAG -3'
6F	5'- GCTGCTATTCGTTTCTAACGTTTAC -3' 5'- GTCACTCCTGCTGGACCTACT -3'
6H	5'- TTGCCCTTCACGGCCACTACC -3' 5'- CCGGGCTTGTCTGCGACTTA -3'
Controllo positivo	5'- GGGAATATAAGGAGCGCACA -3' 5'- TCGGTAAAACGGTCAGGTC -3'
Controllo negativo	5'- CGAGGCGTGAGTTATTCCTG -3' 5'- CTCTTGTGGCTGAGCTCCTT -3'

Tabella 17: oligonucleotidi utilizzati nella MeDip-qPCR. Nella colonna a sinistra sono indicate quali sequenza CpG del promotore di GRP78 sono state prese in analisi. Il controllo positivo si riferisce a regioni di DNA genomico con documentato arricchimento dopo MeDip. Il controllo negativo si riferisce a regioni di DNA genomico non arricchite dopo MeDip.

I motivi che hanno portato ad analizzare solo 7 dei 26 campioni di DNA estratti da tessuto epatico affetto da HCC e solo il 50% delle sequenze CpG sono strettamente economici. Il criterio di esclusione per le sequenze CpG si è basato sulla mancanza di specificità di alcuni oligonucleotidi, in quanto la sequenza a monte di GRP78 è molto conservata.

9. Analisi Statistica

Le frequenze genotipiche sono state calcolate mediante conta genica diretta. La distribuzione di casi e controlli per i polimorfismi rs2304052 T/C, rs4444903 A/G, rs231775 A/G, rs430397 G/A e rs738409 G/C rispetta l'equilibrio di Hardy-Weinberg ($p^2 + 2pq + q^2 = 1$, dove p è la frequenza dell'allele di riferimento e q quella dell'allele di rischio). Le frequenze alleliche e genotipiche di casi e controlli sono state messe a confronto mediante tabelle di contingenza 2X2 e 2X3. La significatività è stata calcolata mediante *p value* e Odds Ratio. La *p value* è stata calcolata mediante test del Chi quadro Pearson. Solo i valori di *p value* < 0,05 sono stati considerati significativi. I modelli statistici utilizzati sono serviti a calcolare l'associazione tra le varianti genetiche in analisi e il rischio di CE nella popolazione siciliana. Sono stati utilizzati oltre al modello classico, anche quello dominante e recessivo.

Per determinare la correlazione tra i valori di espressione del trascritto di GRP78 e lo stato di metilazione del suo promotore è stato calcolato il coefficiente di correlazione di Bravais-Pearson. Tutte le analisi statistiche sono state condotte con il programma Systat 12 for Windows® (Systat Software. Inc.©).

10. Immunoistochimica

Le reazioni di immunoistochimica sono state condotte presso il dipartimento di Anatomia Patologica del Policlinico Universitario di Palermo, grazie al supporto del personale del dipartimento. Ciascuna sezione di tessuto, incluso in paraffina, è stata fissata su vetrini standard e sottoposta a una reazione di immunoistochimica mediante anticorpo specifico per GRP78: Anti-GRP78 BiP antibody (Abcam®). La rivelazione è stata effettuata tramite il cromogeno AEC (Amminoetilcarbrazolo).

L'analisi è stata effettuata sulla base di due criteri: percentuale di cellule positive e intensità della colorazione. Per ognuno dei due criteri è stato attribuito un punteggio che va da 1 a 3. Per il primo criterio: 1 equivale dallo 0% al 25% di cellule positive alla colorazione, 2 dal 26% al 50% di cellule positive, 3 a più del 50% di cellule positive. Per il secondo criterio: 0 equivale ad una colorazione negativa, 1 a un'intensità della colorazione debole, 2 a un'intensità moderata, 3 a un'intensità forte. Sommando questi punteggi si ottengono tre classi: un punteggio totale tra 0 e 2 equivale a un campione negativo, un punteggio tra 3 e 4 a un campione debolmente positivo, un punteggio tra 5 e 6 a un campione positivo.

11. Soluzioni

TBE 5X

Tris base	54g
Acido borico	27,5g
0.5 M EDTA (pH 8.0)	20ml
Portare a volume con 1l di H ₂ O.	

Gel di agarosio per separazione DNA

Agarosio fino a raggiungere la concentrazione desiderata nel volume stabilito

Gel Red®	1µl/30ml TBE1X
----------	----------------

TBE 1X fino al volume stabilito

Gel di agarosio-formaldeide

Agarosio fino a raggiungere la concentrazione desiderata nel volume finale stabilito

TBE 1x fino al volume desiderato

Una volta dissolto l'agarosio, lasciare raffreddare la soluzione fino a circa 60 °C e aggiungere:

Formaldeide al 40% fino ad una concentrazione	2,2M
---	------

MOPS 10X fino ad una concentrazione	1X
-------------------------------------	----

Gel Red®	1µl/30ml TBE1X
----------	----------------

Tampone MOPS 10X

MOPS (acido morfolinpropansulfonico) 0,2M

Sodio acetato 50mN

EDTA 5 mM

Portare a pH 7.0 con NaOH 1M e sterilizzare per 25 minuti a 121 °C in autoclave.

Tampone denaturante per RNA

Per 15 ml di soluzione:

Formamide deionizzata 100%	10ml
Formaldeide 40%	3.5ml
Tampone MOPS 10X	1,5ml

La formamide è deionizzata mescolando 20g di resina a scambio ionico Amberlite MB3 (o MB1) per 15 minuti.

Tampone di corsa per elettroforesi di RNA

Tampone MOPS	1X
Formaldeide	2.2M

Buffer IP 2X

Tampone Sodio-Fosfato, pH 7.0	20mM
NaCl	280mM
Triton X-100	0,1%

Tampone Sodio-Fosfato, pH 7.0

Per 200ml di soluzione:

Sodio fosfato monobasico, monoidrato 0,2M	39ml
Sodio fosfato dibasico 0,2M	61ml
H ₂ O milliQ fino a 200ml	

Tampone di digestione per proteinasi K

SDS (laurilsolfato di sodio)	10%
EDTA	50 mM
Tris-Cl (pH 7.4)	10mM
Proteinasi K	3mg/ml

Discussione

Il carcinoma epatocellulare ha un'incidenza variabile nel mondo, probabilmente a causa della diversa distribuzione dei fattori di rischio e della variabilità genetica tra le diverse razze nel mondo. Oggi, inoltre, è sempre più evidente che l'incidenza ma soprattutto la mortalità del CE dipendono dai piani sanitari e dai protocolli di prevenzione e di diagnosi precoce messi in atto da ciascun paese. Ad Osaka, in Giappone, ad esempio, l'incidenza è diminuita da 41,9 a 24 abitanti su 100.000, dal 1995 al 2003, grazie ad una efficiente politica di prevenzione e una diagnostica di alto livello (Tanaka H., *et al.*, 2008). Il CE è una delle neoplasie per la quale non esistono dei marcatori specifici, che possano coadiuvare il clinico nella diagnosi o durante il decorso della malattia. Il compito non semplice di trovare una molecola del genere è affidato alla ricerca di base. Il CE è una patologia complessa, nella quale sono coinvolte parecchie alterazioni genetiche ed epigenetiche e nella quale, si intrecciano le molteplici vie di segnalazione implicate nella trasformazione maligna degli epatociti e nella progressione tumorale. Un marcatore capace di individuare i soggetti a rischio di CE, al fine di individuare l'eventuale malattia in un momento precoce sarebbe di grande utilità.

I polimorfismi a singolo nucleotide (SNP) possono essere utilizzati come marcatori; tuttavia va ricordato che essi hanno un'ampia variabilità nelle diverse regioni del mondo e quindi la loro validità è strettamente connessa con le frequenze genotipiche e la complessità genetica della popolazione in cui li si studia.

Nel presente studio sono stati scelti alcuni polimorfismi a singolo nucleotide come candidati a diventare dei marcatori di rischio per il CE. Uno degli obiettivi di questo lavoro è stato infatti quello di valutare le frequenze genotipiche e quindi la validità di tali potenziali bio-marcatori a livello locale. Le peculiarità della popolazione siciliana inoltre, sono del tutto uniche. La Sicilia, in virtù delle sue caratteristiche geografiche, potrebbe ospitare degli isolati genetici.

I potenziali bio-marcatori analizzati nella presente ricerca sono stati i seguenti SNP: l'rs2304052 T/C sul gene *SPARC*, l'rs4444903 A/G sul gene *EGF*, l'rs231775 A/G sul gene *CTLA4*, rs430397 G/A sul gene *GRP78* e rs738409 G/C sul gene *PNPLA3*. Il nostro campione di popolazione è stato genotipizzato per questi polimorfismi tramite PCR-RFLP. Il dato di associazione ottenuto per l'rs430397 G/A sul gene *GRP78* è stato successivamente confermato mediante la KASPar genotyping, una tecnica molto più sensibile e specifica, in quanto procede per genotipizzazione diretta. Qualsiasi errore quindi relativo alla fase di raccolta dei dati può essere escluso. Tutti gli SNPs prima citati, sono stati associati al CE in diverse popolazioni da studi recenti. Molti studi riportano associazioni di SNPs con diverse patologie, incluse le neoplasie. L'effetto della variante di rischio a

volte è diretto, come è stato ipotizzato, ad esempio, per l'rs4444903 A/G sul gene *EGF*. Questa variante genetica potrebbe influenzare positivamente i livelli di espressione di EGF (Tanabe KK., *et al.*, 2008). Altri SNP, come quello l'rs2304052 T/C sul gene *SPARC*, che non comporta una variazione aminoacidica, hanno invece degli effetti ancora da chiarire.

Il primo dato, ottenuto dalla presente ricerca, che sottolinea la particolarità della popolazione siciliana, è quello legato alle frequenze genotipiche dell'allele minore (MAF), ottenute. La MAF dell'rs2304052 T/C sul gene *SPARC* dell'8,8% circa la metà dalla MAF globale presente su dbSNP (NCBI), che arriva quasi al 19%. Similmente la MAF ottenuta per l'rs231775 A/G sul gene *CTLA4* del 30,2% è molto bassa rispetto al 45,2%, MAF di dbSNP. Le frequenze genotipiche ottenute per i polimorfismi rs430397 G/A sul gene *GRP78* e rs738409 G/C sul gene *PNPLA3*, invece, non sono così diverse da quelle globali. Entrambe le MAFs riportate su dbSNP, del 12% per l'rs430397 e del 28,4% per l'rs738409 infatti, si avvicinano a quelle rispettivamente del 9,2% e del 30% della nostra distribuzione. Le diverse MAF ottenute testimoniano l'unicità della popolazione esaminata.

L'analisi statistica condotta sul nostro campione di popolazione, per verificare la relazione degli SNPs con il rischio di CE, ha messo in evidenza la non significatività degli indici di associazione per gli SNPs rs2304052 sul gene *SPARC*, rs4444903 sul gene *EGF* e rs231775 sul gene *CTLA4*.

L'rs2304052 sul gene *SPARC* è un SNP che porta ad una variazione nucleotidica, la quale non corrisponde ad una variazione aminoacidica. Il polimorfismo è stato messo in relazione al CE in una popolazione dell'Italia del Nord (Segat L., *et al.*, 2009). La MAF ottenuta nello studio appena citato è molto simile a quella nella nostra distribuzione. La MAF dei casi di CE invece è molto diversa, dato il risultato di non significatività da noi ottenuto. Questa differenza potrebbe essere dovuta alla diversa composizione dei casi di CE. Nella nostra distribuzione infatti, la maggioranza dei casi di CE è HCV positiva (88,53%), mentre nello studio appena citato, il campione di casi di CE è molto più eterogeneo, con una buona fetta di soggetti HBV positivi (41% circa), contro solo il 5% circa dei casi siciliani. Questa considerazione lascia prevedere un'assenza di relazione dell'SNP con il virus dell'epatite C. I ruoli di *SPARC* sono legati alle fasi più avanzate di una neoplasia, *SPARC* è associata infatti, soprattutto alla progressione tumorale e al processo di metastatizzazione (Chlenski A., *et al.*, 2010). Nel CE in particolare c'è un incremento dell'espressione che correla con l'angiogenesi tumorale (Lau CP., *et al.*, 2006). Una relazione con l'HCV quindi potrebbe sembrare quantomeno improbabile. D'altro canto anche una relazione con l'HBV è altrettanto incerta. La nostra distribuzione di casi e controlli rispetta l'equilibrio di Hardy Weinberg, quindi è difficile pensare ad

un errore sistematico nella scelta dei soggetti. Non trattandosi di una distribuzione sufficientemente ampia, nel caso di questo polimorfismo, è possibile invece pensare ad un errore statistico di tipo II. Una volta ottenuto un risultato di non significatività di questo tipo, tuttavia, per motivi squisitamente economici, si è deciso di continuare le nostre analisi su una nuova variante polimorfica.

Anche per l'rs4444903 A/G sul gene *EGF*, non è stata rilevata alcuna significatività nell'associazione con il CE. In una fase intermedia del lavoro, tuttavia, si è riscontrata una significatività che ci ha spinto ad incrementare numericamente il campione di popolazione in analisi per questo polimorfismo. L'rs4444903 si trova in posizione 61 sul 5'UTR del gene *EGF* ed è già stato associato a parecchie neoplasie, come il glioma, il cancro allo stomaco e al polmone. L'associazione con il CE è stata ottenuta da alcuni studi, tra i quali due, indipendenti, su un campione nella popolazione nordamericana e su uno nella popolazione francese, rispettivamente mediante uno studio di prospettiva e uno caso-controllo (Abu Dayyeh BK., *et al.*, 2011, Tanabe KK., *et al.*, 2008). La differenza di metodiche con lo studio americano potrebbe essere un fattore determinante per le differenze nei risultati. Il campione di popolazione americano, inoltre contiene diverse razze, per i neri, ad esempio, lo stesso polimorfismo non rappresenta un fattore di rischio, cosa che invece non vale per i bianchi. La diversa composizione razziale e dunque genetica, in questo caso, potrebbe quindi essere un'altra spiegazione del nostro dato, divergente rispetto a quelli prima citati. Un'altra replica del nostro studio in una popolazione cinese ricalca inoltre il risultato di non associazione (Yuan JM., *et al.*, 2013). Anche in questo caso quindi l'eterogeneità delle composizione genetica del campione di popolazione in analisi gioca un ruolo fondamentale. Anche una recente meta-analisi (Zhong JH., *et al.*, 2012) ha messo in evidenza come l'SNP rs4444903 abbia una difficile correlazione univoca con il CE. Altri studi infatti non sono riusciti a mettere in evidenza la maggiore suscettibilità allo sviluppo dell'HCC in pazienti con genotipo G/G per questo locus.

Sembra che l'rs4444903 influenzi i livelli sierici ed epatici di *EGF* (Abu Dayyeh BK., *et al.*, 2011). I pazienti portatori, in omozigosi, dell'allele G hanno infatti livelli maggiori del fattore di crescita rispetto agli altri genotipi. Evidentemente non sempre livelli di *EGF* circolante sono un fattore di rischio per il cancro, come è stato dimostrato dal presente lavoro. Una conferma di questo dato arriva da una ricerca che associa alti livelli di *EGF* circolante con una migliore prognosi per pazienti affetti da cancro alla mammella (Kim JW., *et al.*, 2013).

L'altro dato di non significatività nell'associazione con il rischio di CE riguarda l'rs231775 A/G sul gene *CTLA4*. Questo SNP è un polimorfismo funzionale posizionato sul primo esone del gene *CTLA-4*, che porta alla sostituzione aminoacidica treonina>alanina sul residuo 17 del peptide leader. L'rs231775 è stato associato ad alcune patologie autoimmunitarie e a diverse neoplasie, come il carcinoma al polmone, alla mammella e alla cervice uterina, oltre che al CE (Sun T., *et al.*, 2009; Hu L., *et al.*, 2010; Gu X., *et al.*, 2010). L'allele A è quello di rischio per le neoplasie, mentre sembra sia protettivo per il rischio di malattie autoimmunitarie (Gu X., *et al.*, 2010). Alcuni saggi funzionali, in modelli *in vitro* ed *ex vivo*, hanno dimostrato che l'allele A porta all'espressione di una proteina con una maggiore affinità con il recettore B7.1, incrementando così l'effetto inibitorio sull'attivazione dei linfociti T e sulla loro proliferazione (Sun T., *et al.*, 2009). L'effetto di questa variante nelle patologie autoimmunitarie è quindi giustificato. Si ipotizza che nel cancro questo SNP sia associato invece con l'abbassamento dell'immunosorveglianza e in particolare nel CE sia correlato con una debole risposta alle infezioni da HCV e HBV.

I nostri risultati smentiscono i dati che associano il rs231775 al CE. Anche in questo caso l'equilibrio di Hardy Weinberg è rispettato nella distribuzione genotipica di questo SNP nel nostro campione di popolazione, per cui un errore nella selezione dei casi di CE è da escludere. Una possibile spiegazione del nostro risultato è data dal fatto che le associazioni fino ad ora presenti in letteratura sono state ottenute su popolazioni asiatiche. Suscita curiosità la MAF da noi ottenuta (30,2%), che è molto lontana infatti dalla MAF del 70% delle etnie caucasiche presente sul sito dell'HapMap Project (<http://hapmap.ncbi.nlm.nih.gov/>) ed è invece molto simile a quella asiatica, del 33,1%. Questo dato singolare potrebbe essere ancora una volta attribuito all'unicità genetica della popolazione siciliana. Un altro aspetto da tenere in considerazione è la differenza nella composizione della coorte di casi di CE presa in analisi dagli studi asiatici e dal presente studio. In uno di questi infatti, l'eziologia è da attribuirsi interamente all'HBV (Gu X., *et al.*, 2010) mentre in un altro la composizione della coorte è piuttosto eterogenea (Gu X., *et al.*, 2010), con una elevata presenza di soggetti coinfectati con HBV e HCV (75%). I casi di CE da noi studiati sono invece per quasi il 90% HCV positivi. L'effetto dell'rs231775 potrebbe essere quindi, nel caso del CE, non specifico per l'HCV e legato invece all'infezione da HBV o all'alcool. Un'ipotesi potrebbe essere legata alla maggiore capacità del virus dell'epatite C di sfuggire al sistema immunitario rispetto al virus dell'epatite B.

La principale significatività statistica ottenuta dalla presente ricerca è relativa all'associazione dell'SNP rs430397 G/A sul gene *GRP78*. La genotipizzazione è stata effettuata sia tramite PCR-RFLP che tramite KASPar™. Anche in questo caso la distribuzione dei genotipi rispetta l'equilibrio di Hardy

Weinberg. Questo esclude ogni errore di selezione dei soggetti. Si potrebbe ipotizzare un errore statistico di tipo II, dato che la nostra coorte è di media grandezza rispetto ai tipici studi caso-controllo. Il dato di associazione però è confermato anche da altri studi in popolazioni molto differenti e geograficamente distanti.

La variante polimorfica rs430397 è localizzata al 3' terminale del quinto introne del gene *GRP78* (Glucose-Regulated Protein, 78kDa). Questo polimorfismo è stato recentemente associato, tramite alcuni studi caso-controllo, nella popolazione cinese, alla cirrosi epatica e al CE (Zhu X., *et al.*, 2009; Zhu X., *et al.*, 2011). Gli autori dei sopracitati studi propongono l'esistenza di un meccanismo tramite il quale l'rs430397 influenzi i livelli di espressione di GRP78. Le dirette implicazioni biologiche dell'rs430397 sono tuttavia ancora sconosciute. GRP78 è uno chaperon molecolare che si attiva a causa di uno stress del reticolo endoplasmatico (RE). Un'altra delle funzioni di GRP78 è quella di sequestrare il Ca²⁺ intracellulare e controllare l'omeostasi del Ca²⁺ nel RE. Si pensa che questa sua attività sostenga la sopravvivenza cellulare (Zhang L-H., *et al.*, 2010). Il ruolo di GRP78 nel cancro è ancora oggetto di dibattito.

Nonostante la MAF ottenuta dagli studi condotti sulla popolazione cinese sull'rs430397 sia intorno al 17%, circa il doppio di quella del nostro campione di popolazione, l'analisi statistica sulle frequenze di casi e controlli ha mostrato dei valori di significatività sia genotipica che allelica. I valori di ORs ottenuti sono in linea con i *p values* di Pearson. In questo caso l'associazione è così forte da andare oltre l'eterogeneità genetica. È quindi ipotizzabile un meccanismo molecolare che possa concretamente aumentare il rischio di CE per i soggetti portatori della variante allelica. Il dato di significatività diventa più marcato analizzando le frequenze genotipiche mediante modello dominante. Questo ci suggerisce un effetto di dominanza dell'allele A su quello G. Questo dato è confermato anche dagli studi sulla popolazione cinese prima citati (Zhu X., *et al.*, 2009; Zhu X., *et al.*, 2011). A supporto di questa ipotesi c'è inoltre la bassissima frequenza di omozigoti A, presente nel nostro campione di popolazione.

Includendo nell'analisi statistica solo i soggetti HCV positivi i dati di associazione si rafforzano sia con il modello classico sia con quello dominante. Questo ci suggerisce una possibile relazione tra l'rs430397 e il virus dell'epatite C. Ci sono delle evidenze che il virus dell'epatite C causi ER stress sia *in vivo* che *in vitro* (Asselah T., *et al.*, 2010). Come è noto lo stress del reticolo endoplasmatico provoca l'attivazione del complesso UPR, il quale partecipa all'amplificazione dei processi infiammatori. L'rs430397 su *GRP78* potrebbe quindi avere un ruolo nell'esacerbare questi processi,

facilitando uno stato di infiammazione epatica, presupposto imprescindibile per l'ambiente pro-tumorale del tessuto epatico. Tramite un meccanismo diverso, anche il virus dell'epatite B causa RE stress (Garg AD., *et al.*, 2012). La proteina pre-S infatti è ritenuta nel RE causando stress e di conseguenza infiammazione. È stato dimostrato che l'rs430397 è associato al CE anche in soggetti HBV positivi (Zhu X., *et al.*, 2011). La relazione con l'HCV da noi ipotizzata trova quindi un'altra conferma.

Il dato di significatività nell'associazione tra l'rs430397 e il CE ci ha indotto a indagare più a fondo tale relazione. Dato che l'SNP su *GRP78* potrebbe avere delle relazioni con il virus dell'epatite C e di conseguenza con l'infiammazione cronica, si è voluto verificare una sua implicazione nella suscettibilità al processo di cirrosi epatica causato. Per questo è stato effettuato un altro studio caso-controllo sull'rs430397 su *GRP78*, che ha incluso soggetti affetti da cirrosi epatica. Non sono stati tuttavia rilevati particolari valori di significatività. In questo caso non è da escludere un errore statistico di tipo II, in quanto la dimensione della coorte di casi di cirrosi epatica è relativamente piccola. Se questo dato fosse però attendibile si potrebbero trarre altre conclusioni. Il polimorfismo su *GRP78* potrebbe rappresentare un fattore di suscettibilità solo in una fase avanzata della malattia epatica e non sarebbe quindi un fattore di rischio per l'infezione di HCV o per una cirrosi causata da altri fattori. Ci sono tanti esempi di alterazioni che rimangono silenti fino a quando non si aggiungono ad altre anomalie, questo è il caso di molte neoplasie o del processo di senescenza cellulare. L'allele di rischio dell'SNP rs430397 potrebbe quindi esibire il suo effetto solo insieme ad altre alterazioni, genetiche o epigenetiche.

Nonostante tutte queste speculazioni, non è ancora chiaro il ruolo diretto del polimorfismo su *GRP78* nella suscettibilità al CE. Per questa ragione nella presente ricerca si è voluta studiare la sua eventuale influenza sul trascritto di *GRP78*. Ci sono infatti delle evidenze che un SNP intronico possa influenzare l'espressione di un gene (Do TN., *et al.*, 2010).

Poiché l'RNA isolato dai tessuti epatici inclusi in paraffina a nostra disposizione è risultato degradato, la nostra attenzione si è spostata sui modelli cellulari. Sono state coltivate alcune linee cellulari umane derivate da CE. Tutte le linee cellulari genotipizzate sono risultate omozigoti per l'allele G. L'espressione del trascritto di *GRP78* è stata inferiore rispetto a quella rilevata in un tessuto epatico sano, per la maggior parte delle linee cellulari analizzate, solo la linea HepG2 ha mostrato dei livelli di espressione di circa il doppio rispetto al tessuto sano. Quindi probabilmente il polimorfismo rs430397 non influenza i livelli di mRNA; il suo effetto va quindi cercato altrove. Il numero di linee

cellulari analizzate non è certo tale da consentire un'analisi statistica ma il trend sembra chiaro. Il dato in questione va sicuramente attestato attraverso ulteriori ricerche.

I livelli di trascritto di *GRP78*, così diversi fra loro, ma anche rispetto ai livelli rilevati sul tessuto epatico sano, hanno destato subito curiosità. È stato quindi analizzato il promotore di *GRP78*. Mediante una ricerca bioinformatica sono state individuate 6 isole CpG. La più grande è sovrapposta al sito di inizio della trascrizione, le altre si trovano una dopo l'altra a monte di questa. Non esistono, allo stato attuale, dati in letteratura che documentano lo stato di metilazione di *GRP78* in un tessuto epatico affetto da CE. È stato portato avanti quindi uno studio sullo stato di metilazione del promotore di *GRP78*. Non avendo potuto utilizzare la tecnica del pirosequenziamento, la migliore per specificità e sensibilità, causa i suoi costi estremamente elevati, lo stato di metilazione del promotore di *GRP78* è stato analizzato con la tecnica MeDip, anch'essa molto affidabile anche se con costi molto più contenuti. Abbiamo analizzato il 50% delle sequenze presenti sulle isole CpG, su un numero limitato di campioni di DNA genomico.

Il dato che è subito risultato evidente è stata la marcata differenza tra lo stato di metilazione sulle linee cellulari e quello rilevato sui tessuti di CE. Le linee HepG2, Hep3b, Huh7 SK-HEP presentano in media uno stato di ipermetilazione. I tessuti di CE mostrano invece un generalizzato stato di ipometilazione. Questa differenza, chiaramente va valutata con estrema prudenza considerando le grandi differenze che esistono tra colture cellulari, cresciute in vitro e tessuti isolati da un sistema in vivo.

I dati sulle linee cellulari umane non hanno mostrato, globalmente, una correlazione tra lo stato di metilazione del promotore di *GRP78* e i suoi livelli di trascritto. È stata invece ottenuta una correlazione inversa escludendo dall'analisi il contributo della linea cellulare HepG2. È possibile quindi ipotizzare che, nonostante lo stato di iper-metilazione del promotore, gli elevati livelli di espressione del trascritto di *GRP78* nella linea HepG2 siano influenzata da altri fattori, nella sfera genomica o proteomica. Per quanto riguarda invece le altre linee cellulari, i bassi livelli di *GRP78*, rispetto al tessuto epatico normale, sono giustificati da uno stato di iper-metilazione del promotore che reprime la trascrizione.

Un altro interessante aspetto è il dato di ipometilazione ottenuto sui tessuti di CE, rispetto all'ipermetilazione del tessuto epatico normale. Se questo dato dovesse essere confermato ad uno stadio precoce o di CE o pre-canceroso, come la cirrosi, si potrebbe considerare lo stato di metilazione del promotore di *GRP78* un marcatore di malattia. Questo potenziale marcatore

epigenetico potrebbe rappresentare un supporto per il clinico nel determinare i soggetti più a rischio o essere correlato ad un particolare tipo istologico di CE. Il CE, infatti è uno dei carcinomi umani nel quale l'alterazione dello stato di metilazione dei promotori è molto frequente (Stefanska B., *et al.*, 2011; Sceusi EL., *et al.*, 2011), inoltre, altri tentativi di individuare marcatori legati allo stato di metilazione di specifici geni sono già stati documentati (Jung N., *et al.*, 2012). Infine il dato di ipometilazione ottenuto potrebbe far riconsiderare, con la dovuta cautela, gli approcci terapeutici che si basano sul trattamento con agenti demetilanti.

Lo stato di ipometilazione del promotore di *GRP78* correla, a grandi linee con il dato dell'analisi immunoistochimica. I tessuti infatti sono tutti positivi per GRP78, e, inoltre, il tessuto che mostra una maggior intensità di segnale è proprio quello che presenta lo stato di ipometilazione più marcato. Da questa correlazione si può dedurre che i livelli di GRP78 in un tessuto epatico affetto da CE non siano bassi e che probabilmente sono influenzati dallo stato di metilazione del promotore di *GRP78*.

In seguito ad una collaborazione con il Royal Free Campus dell'University College (UCL) di Londra abbiamo analizzato l'intera genoteca in nostro possesso anche per l'SNP rs738409 G/C sul gene *PNPLA3*. Questo polimorfismo porta alla variazione aminoacidica isoleucina>metionina in posizione 148 della proteina PNPLA3, un enzima dalla funzione ancora poco chiara, chiamato anche adiponutrina. Questa variante polimorfica è strettamente associata a diverse patologie che colpiscono il fegato come la steatoepatite, la cirrosi alcolica e la NAFLD (Nischalke HD., *et al.*, 2011), mentre la sua relazione con il CE è ancora incerta.

I nostri risultati indicano una debole associazione con il CE. L'analisi statistica delle frequenze alleliche mostra una significatività appena sotto la soglia dello 0,05. Applicando tuttavia il modello dominante si ottiene un'associazione più forte. Ciò sta ad indicare, probabilmente, come per l'SNP su *GRP78*, che l'allele di rischio ha un effetto dominante sull'altro.

L'rs738409 su *PNPLA3* piuttosto, è associato ad altre patologie, non virali, come la steatosi epatica, noti fattori di rischio per la cirrosi epatica e il CE. Le principali ipotesi che riguardano la funzione dell'adiponutrina sono due. Alcuni studi sostengono che abbia un'attività lipasica e altri una transacilasica. In entrambi i casi la sua eventuale associazione con il CE è considerata indiretta, in quanto promuove stati patologici come la steatosi e la cirrosi epatica. Il risultato di associazione da noi ottenuto, seppur debole, tenendo conto che si tratta per la maggior parte di una coorte composta da soggetti HCV positivi, apre a nuove considerazioni. In particolare non si può escludere

una relazione tra questo polimorfismo e il CE HCV correlato. Un ruolo di fattore di rischio diretto per l'rs738409 su *PNPLA3* è stato già teorizzato (Valenti L., *et al.*, 2013). I nostri risultati tendono a confermare questa ipotesi, almeno nella popolazione siciliana.

Conclusioni

Nel corso di questa ricerca è stato avviato un sistema di raccolta di dati genetici, epigenetici, anagrafici e sierologici riguardanti soggetti sani, affetti da CE e da cirrosi epatica, che potrà essere utilizzato per studi di differente natura, incluso quello caso-controllo. Questo tipo di approccio ha consentito l'identificazione di un profilo di rischio per il CE. E' stato evidenziato per la prima volta che i possessori degli alleli A per l'SNP rs430397 sul gene *GRP78* e C per l'SNP rs738409 sul gene *PNPLA3*, in Sicilia, sono soggetti potenzialmente a rischio di insorgenza di CE, in particolare se HCV positivi. Se questo dato dovesse essere confermato da ulteriori ricerche potrebbe essere un'utile strumento per il clinico nel determinare dei particolari profili di suscettibilità al CE.

È stato dimostrato mediante analisi di cellule *in vitro* che l'SNP rs738409 presente al 3' del quinto introne del gene *GRP78* non influenza in maniera significativa i livelli di trascritto di *GRP78* e che invece tali livelli potrebbero dipendere dallo stato di metilazione delle isole CpG sul promotore di *GRP78*. Questi dati escludono uno dei ruoli, ipotizzati per l'rs430397, nello sviluppo del CE e, di conseguenza, restringono il campo per le future ricerche su questo SNP nell'implicazione del CE.

È stata evidenziata per la prima volta una marcata differenza nello stato di metilazione del promotore di *GRP78* tra linee cellulari derivate da CE HepG2, Hep3B, Huh7 e SK-HEP1 e tumori primari al fegato. Questi ultimi presentano uno stato di marcata ipometilazione che, alla luce del dato di positività, ottenuto mediante immunoistochimica, lascia prevedere alti livelli di espressione di *GRP78*.

Infine, per la prima volta è stato dimostrato che soggetti affetti da CE sono caratterizzati da uno stato di ipometilazione sul promotore di *GRP78* nel tessuto neoplastico, a differenza di quanto accade in un tessuto normale. Se questo dato dovesse essere confermato anche in uno stadio molto precoce di malattia potrebbe esser considerato un'altro potenziale biomarcatore di CE.

Indubbiamente altri studi dovranno essere condotti per confermare l'implicazione dei due SNPs individuati nel CE, per verificare quale sia il loro ruolo diretto o indiretto nella patogenesi della malattia e per verificare lo stato di metilazione del promotore di *GRP78* su tessuto epatico di soggetti affetti da CE.

Bibliografia

- Abu Dayyeh BK, Yang M, Fuchs BC, Karl DL, Yamada S, Sninsky JJ, O'Brien TR, Dienstag JL, Tanabe KK, Chung RT; HALT-C Trial Group. A functional polymorphism in the epidermal growth factor gene is associated with risk for hepatocellular carcinoma. *Gastroenterology*. 2011; **141(1):141-9**.
- Aouacheria A, Navratil V, López-Pérez R et al. In silico whole-genome screening for cancer-related single-nucleotide polymorphisms located in human mRNA untranslated regions. *BMC Genomics*. 2007; **8: 2**.
- Aravalli RN, Steer CJ, Cressman EN. Molecular mechanisms of hepatocellular carcinoma. *Hepatology*. 2008; **48(6):2047-63**.
- Asselah T, Bièche I, Mansouri A, Laurendeau I, Cazals-Hatem D, Feldmann G, Bedossa P, Paradis V, Martinot-Peignoux M, Lebre C, Guichard C, Ogier-Denis E, Vidaud M, Tellier Z, Soumelis V, Marcellin P, Moreau R. In vivo hepatic endoplasmic reticulum stress in patients with chronic hepatitis C. *J Pathol*. 2010; **221(3):264-74**.
- Borlak J, Meier T, Halter R, Spanel R, Spanel-Borowski K. Epidermal growth factor-induced hepatocellular carcinoma: gene expression profiles in precursor lesions, early stage and solitary tumours. *Oncogene*. 2005; **24(11):1809-19**.
- Bradshaw AD, Sage EH. SPARC, a matricellular protein that functions in cellular differentiation and tissue response to injury. *J Clin Invest*. 2001; **107(9):1049-54**.
- Chlenski A, Guerrero IJ, Peddinti R, Spitz JA, Leonhardt PT, Yang Q, Tian Y, Salwen HR, Cohn SL. Anti-angiogenic SPARC peptides inhibit progression of neuroblastoma tumors. *Mol Cancer*. 2010; **9:138**.
- Choi JE, Park SH, Kim KM, Lee WK, Kam S, Cha SI, Kim CH, Kang YM, Kim YC, Han SB, Jung TH, Park JY. Polymorphisms in the epidermal growth factor receptor gene and the risk of primary lung cancer: a case-control study. *BMC Cancer*. 2007; **7: 199**.

- Delhaye M, Louis H, Degraef, et al. Relationship between hepatocyte proliferative activity and liver functional reserve in human cirrhosis. *Hepatology*. 1996; **23**:1003.1011.
- Do TN, Ucisik-Akkaya E, Davis CF, Morrison BA, Dorak MT. An intronic polymorphism of IRF4 gene influences gene transcription in vitro and shows a risk association with childhood acute lymphoblastic leukemia in males. *Biochim Biophys Acta*. 2010; **1802**(2):292-300.
- Dragani TA. Risk of HCC: genetic heterogeneity and complex genetics. *J Hepatol*. 2010; **52**(2):252-7.
- Dyer Z, Peltekian K, van Zanten SV. Review article: the changing epidemiology of hepatocellular carcinoma in Canada. *Aliment Pharmacol Ther*. 2005. **22**(1):17-22.
- Edamoto Y, Hara A, Biernat W, Terracciano L, Cathomas G, Riehle HM, Matsuda M, Fujii H, Scoazec JY, Ohgaki H. Alterations of RB1, p53 and Wnt pathways in hepatocellular carcinomas associated with hepatitis C, hepatitis B and alcoholic liver cirrhosis. *Int J Cancer*. 2003; **106**(3):334-41.
- El Serag HB, Rudolph KL. Hepatocellular Carcinoma: epidemiology and Molecular Carcinogenesis. *Gastroenterology*. 2007; **132**:2557-2576.
- Fukunaga-Kalabis M, Herlyn M. Unraveling mysteries of the multifunctional protein SPARC. *J Invest Dermatol*. 2007; **127**(11):2497-8.
- Garg AD, Kaczmarek A, Krysko O, Vandenabeele P, Krysko DV, Agostinis P. ER stress-induced inflammation: does it aid or impede disease progression?. *Trends Mol Med*. 2012; **18**(10):589-98.
- Gu X, Qi P, Zhou F, Ji Q, Wang H, Dou T, Zhao Y, Gao C. 49G. A polymorphism in the cytotoxic T-lymphocyte antigen-4 gene increases susceptibility to hepatitis B-related hepatocellular carcinoma in a male Chinese population. *Human Immunology*. 2010. **71**:83–87.

- Herceg Z, Paliwal A. Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. *Mutat Res.* 2011; **727(3):55-61**.
- Hu L, Liu J, Chen X, Zhang Y, Liu L, Zhu J, Chen J, Shen H, Qiang F, Hu Z. CTLA-4 gene polymorphism _49 A/G contributes to genetic susceptibility to two infection-related cancers—hepatocellular carcinoma and cervical cancer. *Human Immunology.* 2010; **71:888–891**.
- Jeong SW, Jang JY, Chung RT. Hepatitis C virus and hepatocarcinogenesis. *Clinical and Molecular Hepatology.* 2012; **18:347-356**.
- Jung N, Won JK, Kim BH, Suh KS, Jang JJ, Kang GH. Pharmacological unmasking microarray approach-based discovery of novel DNA methylation markers for hepatocellular carcinoma. *J Korean Med Sci.* 2012; **27(6):594-604**.
- Kim JW, Kim JH, Im SA, Lee KH, Kim JS, Kim TY, Han SW, Jeon YK, Oh DY, Kim TY, Park IA. Serum epidermal growth factor is associated with prognosis and hormone receptor status in patients with HER2-positive metastatic breast cancer treated with first-line trastuzumab plus taxane chemotherapy. *Cancer Chemother Pharmacol.* 2013; **72(5):1023-9**.
- Ladeiro Y, Couchy G, Balabaud C, Bioulac-Sage P, Pelletier L, Rebouissou S, Zucman-Rossi J. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. *Hepatology.* 2008; **47(6):1955-63**.
- Lau CP, Poon RT, Cheung ST, Yu WC, Fan ST. SPARC and Hevin expression correlate with tumour angiogenesis in hepatocellular carcinoma. *J. Pathol.* 2006; **210: 459–68**.
- Lévy L, Renard CA, Wei Y, Buendia MA. Genetic alterations and oncogenic pathways in hepatocellular carcinoma. *Ann N Y Acad Sci.* 2002; **963:21-36**.
- Llovet JM, Chen Y, Wurmback E, Roayaie S, Fiel MI, Schwartz M, Thung SN, Khitrov G, Zhang W, Villanueva A, Battiston C, Mazzaferro V, Bruix J, Waxman S, Friedman SL. A molecular

- signature to discriminate dysplastic nodules from early hepatocellular carcinoma in HCV cirrhosis. *Gastroenterology*. 2006; **131(6):1758-67**.
- Nischalke HD, Berger C, Luda C, Berg T, Müller T, Grünhage F, Lammert F, Coenen M, Krämer B, Körner C, Vidovic N, Oldenburg J, Nattermann J, Sauerbruch T, Spengler U. The PNPLA3 rs738409 148M/M genotype is a risk factor for liver cancer in alcoholic cirrhosis but shows no or weak association in hepatitis C cirrhosis. *PLoS One*. 2011; **6(11):e27087**.
 - Peng S, Lü B, Ruan W, Zhu Y, Sheng H, Lai M. Genetic polymorphisms and breast cancer risk: evidence from meta-analyses, pooled analyses, and genome-wide association studies. *Breast Cancer Res Treat*. 2011; **127(2):309-24**.
 - Plentz RR, Park YN, Lechel A, Kim H, Nellessen F, Langkopf BH, Wilkens L, Destro A, Fiamengo B, Manns MP, Roncalli M, Rudolph KL. Telomere shortening and inactivation of cell cycle checkpoints characterize human hepatocarcinogenesis. *Hepatology*. 2007; **45(4):968-76**.
 - Sambrook J, Fritsch EF and Maniatis T. Molecular Cloning- a Laboratory Manual 2nd edn, *Cold Spring Harbor Laboratory press*. New York.
 - Sanyal AJ, Yoon SK, Lencioni R. The etiology of hepatocellular carcinoma and consequences for treatment. *Oncologist*. 2010; **15 Suppl 4:14-22**.
 - Sceusi EL, Loose DS, Wray CJ. Clinical implications of DNA methylation in hepatocellular carcinoma. *HPB (Oxford)*. 2011; **13(6):369-76**.
 - Segat L, Milanese M, Pirulli D, Trevisiol C, Lupo F, Salizzoni M, Amoroso A, Crovella S. Secreted protein acidic and rich in cysteine (SPARC) gene polymorphism association with hepatocellular carcinoma in Italian patients. *J Gastroenterol Hepatol*. 2009; **24(12):1840-6**.
 - Shiraha H, Yamamoto K, Namba M. Human hepatocyte carcinogenesis (review). *Int J Oncol*. 2013; **42(4):1133-8**.

- Sicklick JK, Li YX, Jayaraman A, Kannangai R, Qi Y, Vivekanandan P, Ludlow JW, Owzar K, Chen W, Torbenson MS, Diehl AM. Dysregulation of the Hedgehog pathway in human hepatocarcinogenesis. *Carcinogenesis*. 2006; **27(4):748-57**.
- Soresi M, La Spada E, Giannitrapani L, Campagna E, Di Gesaro V, Granà W, Sandonato L, Brancatelli G, Rotolo G, Affronti A, Messina S, Montalto G. Hepatocellular carcinoma: comparison of two different periods at the same center. *Eur J Intern Med*. 2010; **21(2):127-30**.
- Stefanska B, Huang J, Bhattacharyya B, Suderman M, Hallett M, Han ZG, Szyf M. Definition of the landscape of promoter DNA hypomethylation in liver cancer. *Cancer Res*. 2011; **71(17):5891-903**.
- Sun T, Hu Z, Shen H, Lin D. Genetic Polymorphisms in Cytotoxic T-Lymphocyte Antigen 4 and Cancer: The Dialectical Nature of Subtle Human Immune Dysregulation. *Cancer Research*. 2009; **69:6011-6014**.
- Sur I, Tuupanen S, Whittington T, Aaltonen LA, Taipale J. Lessons from functional analysis of genome-wide association studies. *Cancer Res*. 2013; **73(14):4180-4**.
- Tanabe KK, Lemoine A, Finkelstein DM, Kawasaki H, Fujii T, Chung RT, Lauwers GY, Kulu Y, Muzikansky A, Kuruppu D, Lanuti M, Goodwin JM, Azoulay D, Fuchs BC. Epidermal growth factor gene functional polymorphism and the risk of hepatocellular carcinoma in patients with cirrhosis. *JAMA*. 2008; **299(1):53-60**.
- Tanaka H, Imai Y, Hiramatsu N, Ito Y, Imanaka K, Oshita M, Hijioka T, Katayama K, Yabuuchi I, Yoshihara H, Inoue A, Kato M, Takehara T, Tamura S, Kasahara A, Hayashi N, Tsukuma H. Declining incidence of hepatocellular carcinoma in Osaka, Japan, from 1990 to 2003. *Ann Intern Med*. 2008; **148(11):820-6**.
- Tiegs G, Lohse AW. Immune tolerance: what is unique about the liver. *J Autoimmun*. 2010; **34(1):1-6**.

- Valenti L, Dongiovanni P, Ginanni Corradini S, Burza MA, Romeo S. PNPLA3 I148M variant and hepatocellular carcinoma: a common genetic variant for a rare disease. *Dig Liver Dis*. 2013; **45(8):619-24**.
- Venook AP, Papandreou C, Furuse J, de Guevara LL. The incidence and epidemiology of hepatocellular carcinoma: a global and regional perspective. *Oncologist*. 2010; **15 Suppl 4:5-13**.
- Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schübeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet*. 2005; **37 (8): 853–62**.
- Yang JD, Nakamura I, Roberts LR. The tumor microenvironment in hepatocellular carcinoma: current status and therapeutic targets. *Semin Cancer Biol*. 2011; **21(1):35-43**.
- Yuan JM, Fan Y, Ognjanovic S, Wang R, Van Den Berg D, Govindarajan S, Yu MC. Genetic polymorphisms of epidermal growth factor in relation to risk of hepatocellular carcinoma: two case-control studies. *BMC Gastroenterol*. 2013; **13:32**.
- Zhang LH, Zhang X. Roles of GRP78 in physiology and cancer. *J Cell Biochem*. 2010; **110(6):1299-305**.
- Zhong JH, You XM, Gong WF, Ma L, Zhang Y, Mo QG, Wu LC, Xiao J, Li LQ. Epidermal growth factor gene polymorphism and risk of hepatocellular carcinoma: a meta-analysis. *PLoS One*. 2012; **7(3):e32159**.
- Zhu X, Chen L, Fan W, Lin MC, Tian L, Wang M, Lin S, Wang Z, Zhang J, Wang J, Yao H, Kung H, Li D. An intronic variant in the GRP78, a stress-associated gene, improves prediction for liver cirrhosis in persistent HBV carriers. *PLoS One*. 2011; **6(7):e21997**
- Zhu X, Chen MS, Tian LW, Li DP, Xu PL, Lin MC, Xie D, Kung HF. Single nucleotide polymorphism of rs430397 in the fifth intron of GRP78 gene and clinical relevance of primary

hepatocellular carcinoma in Han Chinese: risk and prognosis. *Int J Cancer*. 2009; **125(6)**:1352-7.

Ringraziamenti

Ringrazio per il loro supporto intellettuale ed economico:

Prof. Mario La Farina

Prof.ssa Michela Giuliano

Prof. Giuseppe Montalto

Dott. Manlio Vinciguerra

Dott. Marcello Cervello

Dott.ssa Lydia Giannitrapani

Dott. Maurizio Soresi

Dott.ssa Rajai Al Jehani

Prof.ssa Marsha Morgan

Prof. Marco Falasca

Dott. Michael Way

Dott.ssa Michelle Holland

Prof. Massimo Pinzani

Dott. Manfredi Rizzo

Un particolare ringraziamento va ai ragazzi del laboratorio CLADIBIOR

Ringrazio anche:

Dott. Giuseppe Nugara

Dott.ssa Monica Failla

Dott.ssa Michela Borgesan

Dott. Fulvio Barone

Dott.ssa Elisa Campagna

Dott. Walter Granà

Dott. Massimiliano Greco

Ma soprattutto:

La mia famiglia, la mia ragazza e i miei amici di sempre.